SYMPOSIUM

Legislating and Litigating Health Care Rights Around the World
GUEST EDITORS COLLEEN M. FLOOD,
LAWRENCE O. GOSTIN AND LANCE GABLE

636 Introduction: Legislating and Litigating Health Care Rights Around the World
Colleen M. Flood, Lance Gable and Lawrence O. Gostin

641 Rights to Specialized Health Care in Norway: A Normative Perspective
Ole Frithjof Norheim

650 Joe Public v. The General Public: The Role of the Courts in Israeli Health Care Policy
Carmel Shalev and David Chinitz

660 Accountability for Rationing – Theory into Practice
Christopher Newdick

669 Just Medicare: The Role of Canadian Courts in Determining Health Care Rights and Access
Colleen M. Flood

681 “Prioritization”: Rationing Health Care in New Zealand
Joanna Manning and Ron Paterson

698 Access to Health Care in the Netherlands: The Influence of (European) Treaty Law
André den Exter

711 Ensuring Reasonable Health: Health Rights, the Judiciary and South African HIV/AIDS Policy
Lisa Forman

725 Health Law 2005: An Agenda
Peter D. Jacobson

Plus, more inside...
Symposium Articles

636
Introduction: Legislating and Litigating Health Care Rights Around the World
Colleen M. Flood, Lance Gable and Lawrence O. Gostin

641
Rights to Specialized Health Care in Norway: A Normative Perspective
Ole Frithjof Norheim
This article examines one procedural attempt to use the rights instruments to advance fair access to health care within the context of a Scandinavian type of welfare system. The Norwegian Patients Rights Act and the political processes leading up to the Act are described and evaluated.

650
Joe Public v. The General Public: The Role of the Courts in Israeli Health Care Policy
Carmel Shalev and David Chinitz
Israel's National Health Insurance Law (1994) has produced a setup in which the courts play a relatively limited role regarding access to health care. However, the emergence of regulatory challenges regarding access may invite a greater future role for the courts.

660
Accountability for Rationing – Theory into Practice
Christopher Newdick
Much has been written on health care resource allocation in theory, but how does theory influence practice? Chris Newdick examines the impact of legal and ethical standards on resource allocation in the UK and comments on their achievements and the challenges that remain.

669
Just Medicare: The Role of Canadian Courts in Determining Health Care Rights and Access
Colleen M. Flood
In the recent case of Chevali, the Supreme Court found a constitutional right to purchase private health insurance in the face of long wait times in public Medicare. This paper argues for a greater focus on administrative law to improve governance and the operation of publicly-funded Medicare for all Canadians and not just those able to buy private insurance.

681
“Prioritization”: Rationing Health Care in New Zealand
Joanna Manning and Ron Paterson
We describe recent innovations in New Zealand in the move from implicit to explicit rationing, and assess two significant legal challenges to rationing decisions. The issue whether rationing decisions are discriminatory is considered, our conclusion being that they are legally defensible, provided certain conditions are met.

698
Access to Health Care in the Netherlands: The Influence of (European) Treaty Law
André den Exter
In the Netherlands, the accessibility of health care services has been guaranteed by both national and international social insurance law. In particular the leverage of Community law in strengthening the patient’s right to healthcare is remarkable. Health care has become inextricably linked with the free movement principles.

711
Ensuring Reasonable Health: Health Rights, the Judiciary and South African HIV/AIDS Policy
Lisa Forman
The South African Constitutional Court’s enforcement of the Constitutional right to access health care services illustrates both the transformative potential of enforcing health claims, and the institutional limits placed on enforcement by the Court itself. This article explores the South African experience, focusing on the impact of the Constitutional right on health policy-making and on legal adjudication itself. Thus relevant jurisprudence is examined in light of national health needs and national health policy, focusing in particular on government policies on HIV/AIDS.
Health Law 2005: An Agenda
Peter D. Jacobson
This article takes a pragmatic look at health law in 2005. Health law doctrine in competition policy and patient safety will influence the primacy of markets or medical professionals. Two intangible health law issues are relations between attorneys and physicians, and understanding how the culture of technology affects legal doctrine.

Introduction: Pain Management in the Emergency Department: Current Landscape and Agenda for Research
Sandra H. Johnson

The Social, Professional, and Legal Framework for the Problem of Pain Management in Emergency Medicine
Sandra H. Johnson

Chronic Pain and Aberrant Drug-Related Behavior in the Emergency Department
Knox H. Todd
Emergency physicians commonly evaluate those with chronic pain and substance abuse disorders; they are more common among emergency department patients than in the general population. Concern regarding coexisting substance use disorders among patients with pain influences treatment; however, the relationship between these two conditions is often murky and limited research is available to guide clinical decision making. At both individual and population levels, a balanced perspective in approaching this difficult issue will best serve patients.

Prescription Opioid Abuse in the Emergency Department
Barth L. Wilsey, Scott M. Fishman, Christine Ogden
Patients in pain who seek opioids in the Emergency Department (ED) may appear to be doctor shopping. But such presentations must never exclusively be taken as evidence of aberrant behavior in a patient seeking to obtain opioids in an ED. Emergency physicians are challenged to reserve judgment even when presented with a patient who seems to have drug-seeking behavior, as this impression may be false. Sometimes, there are no other providers available and patients must attempt to find care in the ED as a "last resort." Since there is rarely a single behavior or event that confirms the diagnosis of addiction in a chronic pain patient, the diagnosis usually remains elusive in the ED setting. In order to help ED physicians find a more tangible method for assessing the potential for opioid misuse in the ED, new screening tools are currently being developed. In addition, legislative initiatives and technological advances in the form of prescription monitoring databases may someday make it feasible for an ED physician to screen patients for evidence of multiple providers.

Clinical Guidelines and Policies: Can they Improve Emergency Department Pain Management?
James Ducharme
This article reviews published guidelines and recommendations related to emergency pain management, identifying barriers to their implementation. Legal concerns toward guideline utilization are presented. Improving the (poor) dissemination of new medical information is also discussed. Methods successfully used in improving pain management in other areas of health care will be reviewed in an effort to improve pain management in emergency medicine.
Independent Articles

791
Phantom Tumors and Hysterical Women: Revising our View of the Schloendorff Case
Paul A. Lombardo
Benjamin Cardozo’s opinion in Schloendorff v. Society of the New York Hospital is often seen to foreshadow the modern doctrine of informed consent. But the 1914 decision may actually have undercut existing understandings of a patient’s right to refuse treatment.

802
Discourse on Embryo Science and Human Cloning in the United States and Great Britain: 1984-2002
Matthew Weed
The United States and Great Britain have radically different policy on embryo experiments of all kinds and only the UK has developed legislation banning attempts at human reproductive cloning. This in spite of the fact that American and British pro-life and pro-research forces have developed cases in support of their views on embryo research and human cloning that are based on the same principles. The close similarity of the ethical cases made for and against embryo science and nuclear transfer experiments in these countries is described. The fact that similar ethical cases did not lead to similar policy demonstrates that ethical discourse alone is insufficient to account for the form that policy takes. New modes for considering ethically complex bioethical issues are offered that account for real world factors such as patients’ growing ability to access cures based on treatments rising from controversial biomedical research.

811
Clinical Trials Registries: A Reform that is Past Due
Jennifer L. Gold and David M. Studdert
In this paper, we support the establishment of a prospective registry for clinical trials as a means of promoting drug safety and scientific integrity. First, we outline why clinical trial registration is important today. Second, we describe registry initiatives by government and private organizations. Third, we outline several important limitations to existing registries. Finally, we propose a set of elements that we believe are integral to a successful clinical trials registry.

821
Motion(less) in Limine
Giles Scofield
Suggesting that it is difficult to determine both whether and how a medical ethicist may offer expert opinion testimony in a manner that can be reconciled with the bioethics literature, the author concedes that there is at least one instance in which such testimony may and should be admitted.

834
Dangerous Times for Medicaid
John V. Jacobi
Short-term Medicaid reform proposals, while painful, are relatively minor. Longer-term proposals are more extensive and dangerous, particularly those that would replace Medicaid’s defined benefits with a defined contribution structure, and would devolve eligibility and benefits design power to states.

Columns

844
Currents in Contemporary Ethics
Stacey A. Tovino

851
The Ethical Health Lawyer
Lance Lightfoot

857
Reviews in Medical Ethics
Tia Powell and Barbara A. Noah

866
Recent Developments in Health Law
The Harvard Law and Health Care Society

870
Index of Articles

872
Calendar of Events
The articles in the *JLME* Fall 2005 issue on the obligation of therapeutic beneficence in clinical research and whether the physician investigator and research volunteer relationship is fiduciary, provided only a cursory overview of the prevailing legal and ethical debate.

In the Litton and Miller article,1 “Distinguishing the Ethics of Clinical Research from the Ethics of Medical Care,” the authors contend that the duty of therapeutic beneficence2 in clinical research is incongruous with moral principles implicit in our accepted societal practices in other contexts. The authors bolster their argument by pointing to the practice of medicine in a military and occupational context, stating that in these settings physicians deviate from the individual loyalty that we expect. They argue that these medical practices reflect the principle that the duties binding physicians should change depending upon the societal role that they fulfill in a particular context; thereby rejecting the principle that physicians are always governed by the medical care ethic. Litton and Miller’s reference to the practice of medicine in a military context brings to mind the events of physicians stationed at Guantanamo Bay and Abu Ghraib and the chilling allegations that physicians had breached their professional ethics and the Geneva Convention by participating in the abusive interrogation of prisoners.3 According to the International Red Cross and other investigative committees, physicians divulged vital medical information to military intelligence personnel, and assisted in the design of interrogation strategies, including sleep deprivation, and other coercive methods tailored to each detainee’s medical conditions. Pentagon officials and military physicians assert that the doctors working at the detention camps did not breach medical ethics because the doctors are acting as combatants, not physicians, when they put their knowledge to use for military ends. “When a doctor participates in interrogation, he’s not functioning as a physician, and the Hippocratic ethic of commitment to patient welfare does not apply.”4 Critical observers point out that in denying their status as physicians, military doctors divert attention from an urgent moral challenge to manage conflict between the medical profession’s therapeutic and social purposes. They contend that the therapeutic mission is the profession’s primary role and the core of physicians’ professional identity. “If this mission and identity are to be preserved, there are some things doctors should not do.”5 I agree.

Furthermore, Litton and Miller’s reference to the practice of occupational medicine as an accepted deviation from the medical care ethic ignores considerable case law on the physician’s fiduciary duty to warn an examinee of a medical condition and to refer the examinee for further evaluation and treatment.6

In the article, “The Clinical Investigator as Fiduciary: Discarding a Misguided Idea,” E. Haavi Morreim offers a limited introduction to fiduciary doctrine and then applies it to the physician investigator—research volunteer relationship. Morreim proposes that in order to understand why clinical investigators are not, and can not be, fiduciaries to research volunteers, we must first understand what a fiduciary is. However, Morreim argues from the premise that the characteristics that determine a fiduciary relationship are universally held. But in fact, fiduciary doctrine varies considerably among common law jurisdictions. For example, normative and legal deconstruction of fiduciary doctrine reveals that in Canada, the Supreme Court has formulated broad guidelines as to when fiduciary duties can arise. The Canadian notion of fiduciary obligation is a considerably more flexible concept not readily applied in American jurisprudence. This point is an important one because it greatly informs and advances the debate among clinical research ethicists, and legal scholars, many of whom are quoted in the Fall 2005 *JLME* articles.

In Canada, fiduciary duties attach to an array of relationships; the case law establishes that a finding of a “special relationship” creates fiduciary duties, especially in a physician-patient relationship. For instance, in *Norberg v. Wynnribe*,8 the Supreme Court of Canada reaffirmed not only that a physician has a fiduciary duty to the patient but also it is the inherent nature of the relationship as defined by the patient that gives rise to the fiduciary duty. Additionally, with respect to the fiduciary obligation of a physician investigator, in the disclosure case, *Halushka v. University of Saskatchewan*,9 the Supreme Court described the relationship between researcher and research volunteers as one of “trust and confidence” and that the physician-investigator’s duty to research subjects was “at least as great, if not greater than the duty owed by an ordinary physician or surgeon to his patient.”10 Also, in the *Gomez* case,11 the court affirmed that biomedical research and medical acts are not in opposition to one another, and that clinical research is an integral part of medicine and is undertaken by physicians whose first professional duty is the protection of the health and well-being of individuals. Similarly, in New Zealand, several notable breach of fiduciary duty cases involving physician investigators have initiated debate on extending the scope of a physician investigator’s legal liability to include breach of confidence, con-
Reconciling diverse fiduciary doctrine is valuable beyond advancing theoretical analysis, it is imperative for the advancement of biological and scientific research. The increase in cross border multi-centered trials and international research in general, demands that ethicists and jurists progress to a discourse that seeks to reconcile conflicts of ethical principles and law and facilitate greater collaboration within the international scientific research community.

Madeline M. Motta, M.S.W., J.D.
Doctoral Candidate, Comparative Law Institute
Faculty of Law, McGill University
Montreal, PQ, Canada
MadelineMMotta@AOL.com

2. Litton and Miller use therapeutic orientation, therapeutic beneficence and medical care ethic interchangeably.
4. Id. at 5.
5. Id. at 5.
8. Norberg v. Wyntib (1992) 92 DLR (4th) 449. A physician who supplied painkillers to a patient he knew was addicted in return for sexual favors was liable for breach of his fiduciary obligation to her.
10. Id. at 445; see also, Frame v. Smith (1987) 42 DLR (4th) 81 at 99, the seminal Canadian case that sets out three general characteristics of relationships in which a fiduciary obligation may be imposed.

Paul Litton and Franklin G. Miller
Reply to Madeline M. Motta

Motta’s criticisms are directed towards a straw man, not our article. Our argument does not, in any way, imply that physicians are free from moral constraints when they are not acting as medical care providers. Without question, physicians have significant ethical obligations in the research setting and when serving other important roles. Indeed, we argued that the importance of medical research and the most scientifically rigorous means of achieving it provide reasons to formulate a moral framework for research that departs from clinical care ethics. However, we also argued at length that such a moral framework appropriate to clinical research must include robust protection of the rights and well-being of participants, albeit without embracing the ethics of medical care.

As an adjunct to our main argument, we pointed out that in addition to conducting research, physicians serve legitimate roles other than medical care provider, such as military medicine. Yet Motta’s insinuation that this point would appear to license physician participation in abusive interrogation practices is entirely erroneous. Just as physicians can violate their moral obligations in research, it is possible for them to violate moral obligations in other settings. That there is a legitimate role for physicians in the military setting does not, in any way, license such abuses of military medicine. One does not need to posit that military physicians are bound strictly by clinical care ethics in order to account for the wrongness of physician contributions to such abusive and degrading behavior. The same holds for occupational medicine and forensic psychiatry in the criminal justice system.

Instead of voicing any real criticism of our actual arguments, Motta’s letter, however, suggests an important issue that deserves systematic attention; namely, how should we think about the duties of physicians in divergent medical roles?

Paul Litton, J.D., Ph.D.
Franklin G. Miller, Ph.D.
Department of Clinical Bioethics
National Institutes of Health

Introduction
Legislating and Litigating Health Care Rights Around the World

Colleen M. Flood, Lance Gable and Lawrence O. Gostin

This special edition of the Journal explores the establishment of health care rights through legislation and litigation in a range of jurisdictions across the globe.

The emphasis in this edition is on rights to publicly-funded health care and the struggle to determine a just allocation of public resources in countries as diverse as New Zealand, Israel, Canada, South Africa, Norway, the United Kingdom, and the United States. Much of the emphasis in this collection is on rationing health care in publicly-funded systems. There are also insights for health care systems with more private funding, such as the United States system, where both private and public payers are increasingly looking to ways to legitimately ration or prioritize access to health care even if it is only for the purpose of expanding profit margins.

The contributors to this volume repeatedly address the difficulty of determining how much money the public sector receives. Carmel Shalev and David Chinitz elegantly reference this point, quoting Paul Simon: “one man’s ceiling is another man’s floor.” A basic problem is that those whose needs are met in the private sector may be less willing through tax dollars or service fees to help fund treatment for those in the public sector who are most in need. On the upside, limited availability of public dollars may lead to more efficient use and rigorous oversight of public dollars. Studies consistently show, for example, that the only real difference in terms between Canada and the United States is the amount of money spent, and that there is little or no overall difference in health care outcomes. More money, particularly more private dollars, doesn’t inevitably lead to overall better health outcomes or even necessarily more consumption of health services. The downside is that at times the public system can be cut too close to the bone, resulting in concerns about waiting for access to treatment and leading to dissatisfaction with the public system and calls for increased privatization.

Concerns about publicly-funded health care have resulted in multiple efforts to establish rights to health care. In some jurisdictions – including Norway, Israel, and the Netherlands – such substantive rights have been established through legislation and, in the case of South Africa, enshrined in the constitution. Other countries, such as New Zealand, have not recognized substantive rights but rather procedural rights: the rights to informed consent, confidentiality, to be treated with dignity, and the like. In the United States, concerns about rationing on the part of private managed care insurers have lead states to implement legislation designed to protect rights of those with private insurance and limit rationing measures on the part of managed care insurers. Concerns about the limits and deficiencies of publicly-funded health care have also resulted in increased litigation, as we see in almost all of the jurisdictions discussed by the various contributors to this journal. For example, in the recent case of Chaoulli v. Quebec (Attorney General) the Canadian Supreme Court ruled that there is a constitutional right to purchase private health insurance...
given unacceptably long wait times for health services in the public sector.

Ole Norheim initiates the discussion in “Rights to Specialized Health Care in Norway: A Normative Perspective,” by examining national legislative efforts to establish rights to health care. Many countries, including Norway, have implemented legal rights to some forms of “prioritized,” “basic,” “appropriate,” or “necessary” health services. Long waiting lists as well as media-reports of patients being denied access to potentially effective but costly services have created pressure on political decision-makers to introduce better tools for securing fair access to priority health services.

The inherent difficulty of specifying health care rights in the form of legal entitlements is that an individual assessment needs to be made of each patient-one size does not fit all. In response to this problem, Norheim discusses how Norway revised its blunter waiting time guarantee legislation in favor of the Patient Rights Act 1999. This legislation provides a more nuanced right to care based on the medical need of the patient, the expected utility of care, and a reasonable relationship between costs and the effectiveness of treatment. The Act also specifies that the right to necessary care does not include experimental or “investigational” treatment. But attaining the right balance between individually-tailored decisions and rights that have “teeth” proved more difficult than first thought. The 1999 Act proved to lack sufficient enforcement power from a patient perspective and from September 1, 2004, every patient in Norway was granted a right to necessary health care to be received within an individually specified time limit. It is the medical doctor in charge that specifies the medically appropriate time limit in each individual case. If the time limit is exceeded, the patient will be provided that service at any hospital in Norway or abroad, paid for by regional health authorities.

The Norwegian experience is illustrative in two important ways. First, it demonstrates the challenge of specifying rights in a way that avoids the essentially contested nature of distributive conflicts. Legal theory tends to assume that the use of professional standards can unambiguously define rights through terms such as “basic,” “appropriate,” or “necessary” services. The second lesson that emerges is that rights to access must be given a specific content through procedures that can be perceived as legitimate, and that innovative methods and mechanisms for achieving this standard are needed.

These two themes are further elaborated upon by Carmel Shalev and David Chinitz in their contribution on the Israeli health care system, “Joe Public v. The General Public: The Role of the Courts in Israeli Health Care Policy.” In terms of the difficulty of specifying rights and the contested nature thereof, they comment that “technical tools for rationing will always have to be complemented by an institutional framework that mediates the political conflicts that inhere to health resource allocation decisions.” As mentioned earlier, the politics surrounding rights to health care revolve around distribution (from the rich to the poor and from the healthy to the sick) and size of income – every dollar spent in the public or private sector equates to income for someone.

In Israel, the 1994 National Health Insurance Law (NHI) establishes a highly specific and detailed list of health services to which every citizen is entitled, as well as a process for updating of this list. Political processes determine which health services are covered under the NHI and the amount of funds allocated for this coverage, resulting in a blend of technocratic and political inputs within a binding legal framework.

As Shalev and Chinitz note, tongue-in-cheek, there are more physicians per capita in Israel than in any other country and an equivalent number of lawyers! Not surprisingly then, the courts have been involved in interpreting challenges to government decisions about the coverage of health services under the NHI. Although clarifying that the list of covered health services is a floor and not a ceiling, courts have nevertheless accepted the need for budgetary restraints and the use of evidence-based medicine to determine benchmarks for public funding. This consensus is, however, being tested as new drugs are brought to market that do not meet the public test of cost-effectiveness but nonetheless have some benefit. This generally raises a problem that the Norwegians also have wrestled with: how to deal with administrative criteria that do not speak to individual cases.

The response in Israel has been for courts not to second-guess decisions to fund or not to fund, but rather to demand that the processes of decision-making are implemented fairly and uniformly. As a result all health insurance funds have adopted consistent processes for determining such claims. The willingness of the courts to uphold coverage decisions made by the funds (albeit with heavy heart for those involved) helps to legitimate the processes of decision-making and improve the trust of Israelis in their health care system.

The United Kingdom has not provided for a list of services that patients are entitled to and the government has shied from adopting centralized decision-making processes that limit health coverage under the National Health Service (NHS). Rather, difficult rationing decisions have been delegated to devolved
agencies called Primary Care Trusts. Devolved decision-makers in the UK must balance colliding pressures. The first set of pressures comes from the courts requiring decision-makers not to be too circumspect in what services are covered and to avoid the use of one-size-fits-all guidelines, which are administratively convenient but often discriminatory. The second source of pressure comes from central guidelines emanating both from the National Institute for Health and Clinical Excellence (NICE) and the central government about access. The strictures of a fixed budget also exert pressure on these decision-makers. But, as Chris Newdick describes in his paper, “Accountability for Rationing – Theory Into Practice,” the courts have proved increasingly willing to review decisions of these devolved bodies—particularly as it applies to the decision-making process. In this regard, the courts mirror a shift in the scholarship in health priority decision-making; away from attempts to articulate the principles guiding the substance of decisions and more to the “accountability for reasonableness” model advanced by Norm Daniels and Jim Sabin, where the focus is upon the processes of decision-making. Newdick concludes that the UK courts have been willing to overturn decisions about coverage that are too blunt and insensitive to the particular clinical needs of individual patients.

Canadian courts have been extremely reticent to be involved in the process of allocation within publicly-funded Medicare. In her article “Just Medicare: The Role of Canadian Courts in Determining Health Care Rights and Access,” Colleen Flood compares this trend in the Canadian health care system unfavorably to the trend in the UK. Part of the problem is the focus on challenges to the Canadian Charter of Human Rights and Freedoms in the Canadian system, with courts in general being reluctant to enter into the policy arena to overturn governmental decision-making or laws. As a consequence, redress in administrative law that could well improve the quality and procedural fairness of decision-making in the public sector has been for too long overlooked. A problem, however, for those who aspire to see a greater role for administrative law, is that much of decision-making within the Canadian system is still “buried” deep within governmental bureaucracies. Thus, it is hard for claimants to identify a particular decision-maker or a particular decision that can be readily subjected to judicial review. Despite the reluctance of the courts to expand the range of services publicly funded or to improve the processes of decision-making through case decisions, the Supreme Court of Canada has in the recent decision of Chaoulli indicated a much greater willingness to wade into the contested waters of balancing public and private financing. In Chaoulli, a bare majority of the court found that laws barring the purchase of private health insurance in the province of Quebec were unconstitutional and violated rights of individuals to life, liberty, and security of the person provided for the Quebec Charter of Rights and Freedoms. This decision has been heavily criticized both for the willingness of the court to intervene in these complex policy choices and for the distributive consequences it will likely have on the Canadian health care system. Some commentators have indicated that the decision signals a shift in Canada to “wealth care” instead of “health care.”

New Zealand, unlike Canada but similarly to Norway, Israel, and the UK has also embraced the notion of transparency in rationing health care and the promulgation of clear guidelines—although Joanna Manning and Ron Paterson note in “Prioritization: Rationing Health Care in New Zealand,” that the New Zealand government does prefer, as many do, the language of “prioritization” rather than the harsher language of “rationing.” But similar to Canada and unlike the other countries discussed so far, New Zealand courts have been reluctant to intervene in prioritization and rationing decisions—particularly those involving rationing at an individual patient level. Partly this is because New Zealand is not a litigious country, as evidenced by the embrace of a no-fault accident compensation scheme which largely precludes individuals from bringing personal injury lawsuits for medical malpractice or other health consequences. Similarly, this situation may be attributed to the fact that such rationing decisions are often cloaked as clinical decisions even although imbedded within them are arguably aspects of age and disability discrimination, as Manning and Paterson eloquently describe in the context of the Shortland decision. Another factor may be that there is not a clear expression of rights to non-discriminatory treatment in New Zealand’s Bill of Rights. Thus, provided there is an objective, evidential basis to justify exclusion from treatment on the grounds of age or disability, then such exclusion is legally justifiable in New Zealand.

Manning and Paterson conclude by noting that it is important for a government to be able to justify apparently discriminatory action in rationing health care in order to ensure the sustainability of public health care. The issue of balance between access and sustainability is also picked up by Andre den Exter in “Access to Health Care in the Netherlands: The Influence of (European) Treaty Law,” discussing the Dutch health care system and the impact of decisions by the European Court of Justice (ECJ) on access claims. The ECJ has concluded that health care is a
service in terms of the European Treaty and thus rights to health care provided for in domestic legislation have been inextricably linked with free movement principles. This conclusion enables claims on the part of European citizens for reimbursement for health services in other jurisdictions within Europe, resources available to the state for public health care and an enduring hesitancy on the part of courts to affirmatively review governmental policy-making.

Finally, Peter Jacobson examines the state of the United States health care system, whose population of 45 million uninsured stands in stark contrast to the constitutional commitment in South Africa to positive health care rights. In “Health Law 2005: An Agenda,” Jacobson sets out in broad strokes a new agenda for health law in the US that moves beyond the old debates around tort law. Unlike many of the other countries cited in this issue, the US system relies on a more privatized health care model. Further, there is little legislation providing for rights to health care and little litigation upholding such rights.

Despite the exceedingly different structure in the US compared with the other countries already examined, Jacobson picks up on many of the same themes. For example, he focuses on the issue of governance and inquires about the role of courts to review governance decisions that mediate the balance between access and sustainability (or in the US market-system “margins”). The concern of how to ensure good governance on the part of decision-makers cuts across public insurance, private insurance, for-profit and not-for-profit entities, and even down to individual physicians who, particularly a market-oriented system, may have financial incentives (e.g., gain sharing) to make trade-offs between cost, quality, and access. Jacobson also addresses another issue facing all health care systems regardless of the public/private mix, namely the emergence of new technologies, the desire of patients to have any and all innovations that may provide some marginal benefit however expensive, and the related desire of physicians to satisfy those demands. This, as discussed by other contributors, poses a particularly acute problem for public insurers who must provide benefits from a much more reduced budget. It also raises issues within a private-pay system where the ethical appropriateness of repeated and invasive interventions, particularly at the end-of-life, demands greater consideration. Jacobson’s tour-de-force survey of emerging issues in the US also touches on the issue of patient safety. Every country in the world – apart from the US – seemingly is more persuaded by the logic of the argument in health care that more care is

The liberalization of health care through free trade agreements coupled with domestic rights to health care has undoubtedly improved access for some individuals. These developments, however, do raise questions about the overall sustainability of public health care including in circumstances when individuals seek health services in another jurisdiction due to concerns about timeliness of treatment or constraints on new and innovative technologies in their own jurisdiction. The liberalization of health care through free trade agreements coupled with domestic rights to health care has undoubtedly improved access for some individuals. These developments, however, do raise questions about the overall sustainability of public health care (or in Europe, social insurance).

Many international human rights agreements provide for positive rights to health and health care. And indeed, a recent study shows that over two thirds of all national constitutions have provisions regarding health and health care. Despite this recognition of rights to health in foundational legal instruments, as Lisa Forman discusses in “Ensuring Reasonable Health: Health Rights, the Judiciary and HIV/AIDS Policy,” there has been little uptake of the notion of a positive right to health care by courts. In her chapter she tells the story of both the establishment and interpretation of a positive right to health care in the South African constitution. From the ashes of apartheid rose a constitution that sought to heal in some measure the enormous inequality between black and white South Africans. It was thought that protection of negative rights only – preventing government from interfering with individual autonomy – would result in enduring inequality. Thus, the new South African constitution provides for the progressive realization of positive rights to housing, water, food, social security, and health. The constitutional right to health, as interpreted by the Constitutional Court of South Africa proved powerful enough to cut through governmental intransigence on HIV/AIDS and forced the government to provide for nevirapine – a cheap and effective measure to reduce the incidence of mother-to-child transmission of HIV. The right to health care as interpreted to date by the South African Constitutional court is a qualified one; qualified by the limited
not necessarily better. Perhaps it will be the issue of patient safety – and the startling data on level of iatrogenic injury – that will finally start to shift the mindset of US citizens and decision-makers in this regard.

This collection of essays illustrates the breadth of knowledge to be gained from examining approaches and experiences across jurisdictions. The respective roles of government and the courts play out differently in different contexts. However, the potential of the courts to improve fairness within a system and to reduce inequalities through interpretation of existing legislation manifestly exists in all of these jurisdictions. Their will to do so may depend upon the extent to which they can be persuaded of the legitimacy of their role in this regard. Thus, the examples of courts taking on this role in other jurisdictions may prove more than just of academic interest.

References
Rights to Specialized Health Care in Norway: 
A Normative Perspective

Ole Frithjof Norheim

It is possible to use the courts – or rights instruments – to advance fair access to health care? This article examines this question within the context of the Norwegian public health care system – one special example of the Scandinavian welfare system. In particular, it asks four basic questions: What are the normative justifications for rights to health care? What were the political processes and concerns leading up to the current Patients Rights Act in Norway? What kind of legal status do these rights have? How can rights to access be implemented?

Patient rights do not only concern the right to access to health care; they also include the right to information, the right to participate in decision-making, and informed consent. This article examines only the former aspect, the use of the legal system to secure access to prioritized specialized health care services.

Before turning to these questions, a brief presentation of the context of the Norwegian health care system is needed. Health care systems are fairly similar in the Nordic countries. The exact details vary, but in all countries the systems have been almost exclusively publicly funded through taxation and some co-payment. Most hospitals are publicly owned and managed, although in the last four years more private hospitals have been encouraged. In Norway, the government owns most hospitals, but the provision of services in hospitals and through specialists is a responsibility of regional health authorities.

Financing was traditionally controlled through global budgets that are broken down into hospital budgets and then further divided among various hospital departments. Since 1997 a percentage (40-60%) of the expenses are reimbursed according to the number and kind of patients treated (called activity-based financing). Norway has a fairly strong primary care sector where family physicians, to various degrees, act as gatekeepers to specialist services. All citizens are granted free accommodation and treatment, including medicines, in hospitals. The Act on Municipal Health Care of 1984 obligates municipalities to provide necessary primary care to all of their citizens. Until 1999, there was no corresponding right to specialized health services (secondary and tertiary care).

What are the Normative Justifications for Rights to Health Care?
The first justification for a right to health care is grounded in human rights. Human rights include, but

Ole Frithjof Norheim, M.D., is a physician and a professor in medical ethics. He has also acted as Senior Adviser to The Western Health Region and Helse Bergen, Haukeland University Hospital.
Health care systems are fairly similar in the Nordic countries. The exact details vary, but in all countries the systems have been almost exclusively publicly funded through taxation and some co-payment.

therefore protect equality of opportunity at each stage of the life. From fair equality of opportunity follows that there should be no: “financial or other discriminatory barriers to access to the level of health care that best promotes normal functioning and that best protects the range of opportunities open to individuals in that society.”

On this account, the distinction between services that society should and should not provide is drawn by appealing to the concept of the normal array of opportunities existing in that society. Services directed to restoring, rehabilitating, or sustaining restricted opportunities will fall under the basic tier of care for which society is responsible. The argument is based on the view that health-related opportunities are substantive goods, or all-purpose goods, necessary for functioning in a society. Health benefits are neutral, so to speak, as they are substantive goods or means to develop and pursue a rational life-plan. Access to these substantial goods should be equal. This argument underlies the principle of equal access to health care for everyone. In Daniels’ terms, this takes the form of a legitimate ethical claim, but not necessarily a legal right. The institutionalisation of ethical claims as legal rights might depend on the particular health care system in question. A third possible ethical justification for a right to health care, although not widely advocated, is the instrumentalist argument. One might argue, for example, that utility would be maximized by institutional-
izing a set of rules, e.g. rights, as a coordinating mechanism for constraining individual decisions in order to maximize societal aggregate utility or “as a way of parceling out valued forms of discretion over which individuals are in conflict.”

Rights to health and health care can therefore be justified by a variety of religions, beliefs, and ethical theories, although the content of such rights will vary accordingly. Leaving aside the special case of human rights, however, most ethical distributive theories of health and health care acknowledge that there might be a conflict between, for example, the goal of securing equal access to health related opportunities and the goal of maximizing health related utility. The distribution of scarce health care resources should satisfy both goals: fairness and allocative efficiency.

The goal of allocative efficiency is based on the notion that health resources should be allocated across interventions and population groups to generate the highest possible overall level of population health. Cost-effectiveness information is “crucial in defining a mix of interventions that would be health maximizing in the absence of any constraints on possible decisions except a finite budget.”

The goal of fairness requires that health benefits should be distributed among groups of people in a way that is considered just. Two distributive concerns in particular might be in conflict with efficiency. Empirical studies have shown that people seem to value health benefits for the very sick more than benefits to less ill patients. The argument justifying this concern is indirectly egalitarian. It can be grounded in the principle of equal health (however health is defined). Derek Parfit calls the priority view non-relational egalitarianism. Another aspect of fairness is equality in the distribution of health benefits, where one possible formulation is a principle of equal healthy life expectancy: Each citizen has a legitimate claim to the same healthy life expectancy as everyone else enjoys – within the constraints of what is medically possible. Put in other words, life expectancy in terms of healthy life years should be equalized.

Of course, these concerns for fairness (severity and inequality) must be balanced against other concerns. For some distributive problems, equality can require very large sacrifices in terms of the aggregated sum of health benefits. In the ethical literature of health care distribution, there is less agreement on how to balance these conflicting concerns. From this follows also a certain reluctance to propose substantive health care rights as a way of resolving distributive conflicts under resource scarcity.

The indeterminacy of substantive ethical justifications for the content of rights to health care should not, however, lead us to overlook widely agreed norms of non-discrimination. There is general consensus that individual characteristics such as race, ethnicity, place of origin, religion, age, sex, social status, sexual orientation, and physical or mental disability should be irrelevant in devising rationing criteria. The norm of non-discrimination forms the basis of the principle of equal treatment before the law, or what health economists call “horizontal” equality: equal treatment of equals. The deep disagreements in ethical theory (and the practice of priority setting) concerns disagreement on “vertical” equality: unequal treatment of unequals. Who should be assigned a right to prioritized access to health care: the patient in a very severe condition who can expect a small health benefit at a very high cost (typically cancer patients facing life-or-death decisions), or the less severely ill patient who can gain a larger medical benefit at a lower cost (e.g. patients at risk of cardiovascular disease, where preventive measures are available)?

In summary, ethical theories can and do justify a fundamental right to health care. But assuming scarcity of resources within any health care system, it is harder to specify rights in a way that avoid the essentially contested nature of distributive conflicts. Despite these theoretical worries, many countries, including Norway, have implemented legal rights to some forms of “prioritized,” “basic,” “appropriate,” or “necessary” services. The Norwegian example illustrates one innovative way of resolving these conflicts.

The Political Processes behind Norway’s Patients Rights Act

Legitimate concerns and goals for any reasonable health policy include the need to control escalating costs, achieve efficiency, and secure fair access and distribution of health care services. In broad terms, legitimate governance is based on at least four kinds of “instruments”:

- Economic instruments: budgeting and financing mechanisms including incentive mechanisms, etc.
- Organizational instruments: centralized versus decentralized governance, health sector reform, public or private provision, etc.
- Legal and regulatory instruments: Certification, accreditation, patient rights, etc.
- Educational and professional instruments: professional standards, clinical guidelines, evidence based medicine, motivation and cultural change through education and information campaigns.

In the context of distribution and access to care, some aspects of Norwegian health sector reforms in the last
decade can be seen as a strengthening of legal instruments, perhaps as a counterweight against forceful economic influences, such as incentive-based financing systems. In addition, long waiting lists, as well as media-reports of individual patients being denied access to potentially effective but costly services, have created a pressure on political decision-makers to introduce better tools for securing fair access to priority services.\textsuperscript{18}

The development and implementation of patient rights in Norway has a special history that should be seen against the backdrop of the governmental response to these pressures. Waiting lists have a special place in this history. In 1987 Norway was the first western country to develop national guidelines for priority setting in health care.\textsuperscript{19} The Committee recommended that severity of disease should be the most important criterion for determining priorities, and it identified five separate priority levels. The report suggested that services in the two lowest priority groups should not be publicly funded before a satisfactory level of provision was achieved for the three highest groups. The details of this report are described elsewhere.\textsuperscript{20} The emphasis on concerns for the worst off (the most severely ill) is probably the report’s main characteristic. Its guiding principle was expressed almost in Rawlsian terms: “changes that improve the health status for the better off groups should involve at least the same degree of improvement to the worse off groups.”\textsuperscript{21} The commission in particular emphasized the need to improve psychiatric services, rehabilitation, long term care, and care for people with chronic illness.

The most visible consequence of the Norwegian guidelines was the so-called waiting list guarantee, which passed Parliament in 1990. A guaranteed maximum waiting time of six months was based on the definition of priority group two: non-acute conditions for which delayed treatment would have catastrophic or severe consequences. The system has been evaluated in several studies.\textsuperscript{22} The high number of violations of the guarantee has been of constant political concern. Many clinicians were against the regulations. Kristoffersen and Piene, for instance, found that doctors interpret severity in different ways, and that they are willing to reinterpret their patient’s disease state if it will improve the patient’s priority.\textsuperscript{23}

In 1996 a new commission was asked to revise the national guidelines of 1987. This second commission, also chaired by Professor Inge Lønning, had as its stated goal to recommend changes that could involve clinicians’ day-to-day experience with limit setting decisions. The so-called Lønning II-commission recommended that changes to the current system of prioritisation should take place \textit{ab infra} – from the bottom up.\textsuperscript{24} It proposed that speciality specific working groups are established to make recommendations regarding priority setting within their own specialist fields and in accordance with pre-defined criteria of four priority groups: core services, supplementary services, low priority services, and services which should not be financed or reimbursed by the public health care system. A fifth priority group of “investigative treatment” was also suggested. The aim of inviting advice on clinical priorities was to improve interaction between the political and clinical levels of decision-making through processes that start at the bottom and go upwards before returning to health personnel with day-to-day responsibility for using these guidelines. The commission argued that the reliance on procedures at several levels of the decision-making system is important “because principles and criteria do not in themselves provide solutions to difficult rationing decisions. There is disagreement on the interpretation of evidence, and there are uncertainties and disagreement over which principles and criteria apply.”\textsuperscript{25} The underlying idea was that “a combination of guiding principles, fair procedures and appropriate medical judgements can secure the necessary legitimacy of the decisions made.”

The Lønning II commission also considered using rights to secure prioritized access to care. The most important consideration was that if rights implementation was to be effective, such rights must be clearly defined and easily understood. A right to primary and secondary education is a case in point (every person within a given age group can be assigned this right). “Prioritized care” or “core services,” on the other hand, is hard to define because it depends on factors varying to a large degree within each diagnostic group, i.e., on the condition of the patient in question (how severely ill he or she is) and whether certain medical selection criteria are fulfilled (indicating prognosis and the expected effectiveness). Moreover, there might be some disagreement on the actual interpretation and application of the underlying priority principles. The majority therefore felt that the scope for individual discretion in deciding who should have priority access would necessarily be wide. The complex nature of the issue at hand therefore made it unsuitable for inclusion within legal rights instruments. The minority of the commission argued, on the other hand, that rights – and the use of the legal system – would strengthen the legitimate claims of individual patients for priority services. They also argued that violations of rights should be followed by economic sanctions for the service provider, securing a right that would be binding and not merely a well-intentioned political document.

The Patients’ Rights Act of 1999 followed up on some of the recommendations from the Lønning II commis-
sion, in particular by specifying the criteria for assigning rights to specialized services.26 The Act replaced the former waiting time guarantee. The objective of the new legislation is to ensure the population equal access to health care of good quality by granting patients rights to “necessary” health care. The right to health care is defined according to three criteria:

The patient has a right to necessary specialized health care according to the Patients Rights Act, § 2-1, when:
1. The patient has a condition with reduced prognosis related to life expectancy or quality of life if health care is delayed, and
2. The patient has an expected effect of the health care, and
3. There is a reasonable relation between costs and the effectiveness of the service.27

In addition, there is a requirement of sufficient quality of evidence concerning criteria (1) – (3). The act emphasizes that “the health care provided shall be proven to be effective. There must exist good medical evidence from Norwegian or International studies that the condition can be improved by medical treatment. The right to necessary care does not include experimental or ‘investigational’ treatment.”28 The waiting time guarantee was replaced by the formulation that services shall be provided within time limits judged as medically appropriate by the individual medical provider from case to case. No sanctions were linked to violations of the Act.

The Parliament evaluated the execution of the Act of 1999 in 2003, and concluded that regional health authorities and medical providers had not sufficiently implemented the Act. In a further revision, the Parliament introduced economic sanctions in cases where the right to necessary care was violated. Beginning in September 1, 2004, every patient granted a right to necessary health care shall receive services within a medically appropriate time limit determined by the medical doctor in charge in each individual case. If the time limit is exceeded, the patient will have a right to be provided that service at any hospital in Norway or abroad.29 A central coordinating office helps patients to find an appropriate provider (and negotiates prices with a given number of hospitals, mainly private hospitals in Norway and the other Nordic countries). The regional health authorities must reimburse the expenses.

**Rights to Specialized Health Care under Norwegian Law**

In standard ethical theory, following Dworkin’s definition in his book “Taking Rights Seriously,” individual rights are seen as absolute constraints on other concerns, such as the goal of maximizing aggregate utility in society.30 In modern health care systems, with so much concern for allocating resources according to what is most cost-effective, one might therefore assume that the individual right to health care should be interpreted as a constraint on this goal. A severely ill patient in need of effective treatment would, according to this standard interpretation, be assigned a right to “necessary” care, even if it were very costly. Observe, however, the third paragraph in § 2.1 of the Norwegian Act that incorporates a consideration of cost-effectiveness in the specification of the right itself: “There is a reasonable relation between costs and the effectiveness of the service.”

This unique specification could be seen as making the right to health care a special kind of hybrid: a right to “rationed services.”31 When concerns for the aggregate sum of benefit are incorporated in the right itself, so to speak, the right does not secure an individual’s claims against the interests of the society.32 On the other hand, it could also be seen as an innovative attempt at resolving the inherent ethical conflict outlined in the previous section of this article. It stands out as an alternative to the definition of a right to “the highest attainable standard of health” as well as Dworkins’ definition of rights as trumps.

However, there is a concern that the assignment of rights could be influenced by the actual capacity of the provider in question. This is not unlikely to happen, and there are no mechanisms in the present procedure for the assignment of rights that prevent this outcome.

Moreover, it is worth noticing that there is little tradition in the Scandinavian welfare systems to use the courts to advance access to specialized health care. Patient rights acts have typically been seen as specifying obligations for the providers to secure access to necessary care. Until now, to our knowledge, no cases have been taken to court with reference to the Act of 1999.33 It is likely that the courts will – when the Act has been in function over time – increasingly see cases where patients appeal when not being assigned a right to necessary care. In a somewhat related set of cases, that of appeals concerning reimbursement for expensive treatment abroad, there have been a few court cases already.34 There is no question that the courts can be used to protect the right to care, but it is too early to assess how effective this protection will be.

**Implementing Rights to Specialized Health Care Services: A Continuous Social Learning Process**

How can rights to access be implemented? After the legislative changes in 2004, which introduced economic sanctions, three out of five health regions in Nor-
way have actively started implementing the new law. These implementation efforts assume that right to health care can only be specified through a process involving the clinical speciality groups that are delivering such care. One approach has been the development of clinical recommendations that outline what groups of patients should be assigned a right to necessary services, including advice on appropriate maximum time limits. The resulting rough recommendations establish what groups of patients should be assigned a right to necessary services, but do not satisfy the requirements of enacting more detailed clinical guidelines developed for a particular condition or service.

In the Western region, recommendations have been developed according to the guidance of the Lønning II commission and the requirements set by the “Accountability for Reasonableness” ethical framework. The key question motivating the Accountability for Reasonableness framework is: “Under what conditions should society grant authority to individuals or institutions to set limits to health care?” The answer to this question is that it is possible to specify some characteristics of fair procedures that all who are affected by such decisions can agree are relevant and acceptable. The framework suggests four conditions that should be met for fair procedures:

1. **Publicity Condition**: Decisions regarding both direct and indirect limits to care and their rationales must be publicly accessible.
2. **Relevance Condition**: The rationales for limit-setting decisions should aim to provide a reasonable explanation of how the organization seeks to provide “value for money” in meeting the varied health needs of a defined population under reasonable resource constraints. Specifically, a rationale will be reasonable if it appeals to evidence, reasons, and principles that are accepted as relevant by fair-minded people who are disposed to finding mutually justifiable terms of cooperation.
3. **Revision and Appeals Condition**: There must be mechanisms for challenge and dispute resolution regarding limit setting decisions, and, more broadly, opportunities for revision and improvements of policies in the light of new evidence or arguments.
4. **Regulative Condition**: There is either voluntary or public regulation of the process to ensure that conditions 1-3 are met.

What is unique in the case of Norway is that the framework is tested within the special context of assigning rights to patients. This means that the reasons for limit setting that can be given are constrained by the criteria given by the Patients Rights Act.

The results from the western region so far show that more than 20 such recommendations have been developed for various specialities. The recommendations are of good quality, but have varying levels of precision or accuracy. Examples are given in table 1.

Table 1 is developed by identifying so-called threshold patients within each specialty. A threshold patient is defined as the typical patient ranked lowest in each recommendation, but still given a right to necessary care. Medical needs that qualify for the assignment of rights vary considerably across specialty groups (table 1). For example, in ophthalmology, stenosis of the nasolacrimal duct with tendency for infections, qualifies as a right. In adult psychiatry, by comparison, the threshold patient with anxiety is one who has several episodes of anxiety, with tendency to isolation/avoidance, where anti-depressants have not improved the situation, and where there is high risk for deterioration of the condition if treatment is not provided. There seems to be a much higher threshold for the right to necessary care in the context of adult psychiatry than in the context of ophthalmology. Other, even more strik-
ing examples can be found. The Western health region is now in the process of harmonising the recommendations across patient groups and specialities. The next steps will include, first that each speciality will be asked to justify their chosen threshold by reference to:

1. Severity of disease
2. Effectiveness of the technology
3. Reasonable and acceptable cost-effectiveness
4. The quality of evidence for (1) – (3).

The second step is to conduct a hearing among users, i.e. in the hospital user’s board (representatives of patients and their interest’s organizations). This involves also “translation” of the recommendations to non-technical language. Later, the recommendations, with the given thresholds, will be posted on the internet and made available for patients and general practitioners who refer these patients.

It remains to be seen if the procedure chosen in the Western health region works and can be considered acceptable. So far, experience shows that limit-setting decisions are contested, that it is more difficult to define rights for certain complex types of health care such as drug dependency, psychiatric, and chronic disorders, and that there is a lack of enthusiasm from some professionals.

For the country as a whole, it is too early to evaluate the effect of the recent changes in the Act, but two results can be noted. First, there has for several years been a downward trend in waiting times. This trend continues for all health regions (figure 1).

The trend is probably a result of massive increases in health care expenditures in specialized health care services combined with health sector reforms including patient rights. There are also indications that the Act itself, so far, has little impact. Waiting time for patients assigned a right is shorter than for those not assigned a right, but the difference is only five days for somatic disorders, and 18 days for psychiatric disorders.

The second result to be noted is the wide geographical variations in the percentage of patients being assigned a right. In the Southern region 50% are assigned a right, while in the Northern region, 90% are assigned a right (figure 2).

Variations in capacity and demand cannot fully explain this picture. One possible explanation is that hospitals and health regions interpret and implement the Act and its guidance differently. As mentioned above, there is wide room for discretion, and this may partly explain the difference. If this trend continues, however, high geographical inequality in who are and how many are assigned a right might indicate that the system is not working adequately.

The unique and positive lesson learnt from the Norwegian experience is that it is possible to define the content of rights through a process involving advice from the professions. However, the recommendations developed leave much room for clinical discretion, resulting in large variations between clinical specialties and geographic regions. Open explicit recommendations are therefore only a starting point for a process towards fair and legitimate assignments of right to health care. If it turns out that clinical discretion is exacerbating geographic and specialty inequalities, the whole approach must be revised.
Concluding Remarks: Is it Possible to use the Courts – or the Rights Instrument – to Advance Access to Health Care?

Based on the experience from Norway, it is to early to conclude that it is possible to use rights to advance access to high priority health care. Without a thorough evaluation, we do not know whether it works.

But the Norwegian experience is unique in two important ways. First, it shows that it is hard to specify rights in a way that avoid the essentially contested nature of distributive conflicts. From an ethical perspective, legal theory tends to assume that the use of professional standards can unambiguously define rights in terms such as “basic,” “appropriate,” or “necessary” services. One lesson that emerges is that rights to access must be given a specific content through procedures that can be perceived as legitimate, and that innovative methods and mechanisms for achieving this are needed. The Western Health Region’s attempt to use the ethical framework of Daniels and Sabin, is just one possible approach. Second, from a normative perspective it is often argued that the use of rights is not an acceptable way of resolving distributive conflicts e.g. between individual interests and societal welfare. The Norwegian Act is unique, in the sense that it’s formulation incorporates a consideration for reasonable cost-effectiveness, and thus allows for a balancing between individual’s claims and the interests of the society. More analysis and research is required to examine whether this approach is normatively acceptable or not.

References

4. Id.


25. Id.

26. Id.

27. Id.

28. Id.

29. Id.

30. Id.

31. Id.

32. Id.

33. Id.

34. Id.

35. Id.

36. Id.

37. Id.

38. Id.

39. Id.

“One man’s ceiling is another man’s floor”
Paul Simon

The words of Paul Simon capture the essence of what the courts are called upon to deal with when adjudicating matters of health. Wealthier and healthier neighbors living in the upstairs apartment are, all things being equal, not overly interested in raising the floor of their apartment in order to create more space for those in the apartment below. The floor is tangible, measurable and moveable, and thus a subject where science can contribute much. However, where to locate the floor is a matter of values and thus largely in the realm of politics.

Resorting to the courts in matters of public policy can be seen as the residual of the inability of the remaining two branches of government, the legislature and the executive, to succeed in allocating resources and regulating the system. When the combination of science (bureaucracy) and politics (legislation) falters, enter the judiciary.

Of course health and health care are not the only policy area in which this occurs. Nonetheless, health is, arguably, the area in which decision-makers are most reluctant to accept clear lines of accountability. Health care policy decisions are often perceived as dealing with matters of life and death. Stakeholders seek to avoid being at the place where the buck stops.

At the same time, the last two decades of health reform have increasingly pushed the issue of explicit health care rationing and prioritization to the forefront as the desire to restrain costs clashes with increasing needs and the availability of new health care interventions. At the macro, meso, and micro levels methods such as clinical guidelines, cost-benefit analysis, financial incentives, and priority setting by expert panels manifest the position that not all treatments that might be of benefit can be funded. Recently, and not surprisingly in the United States, the argument has been made that the focus should be on quality, i.e. ensuring that provided services enhance health, and not on cost control. But this begs the question of the process by which “quality” will be determined. One impetus for the appeal to quality might be the fact that United States Supreme Court decisions in the last few

Carmel Shalev, J.S.D., teaches health and human rights law at the Tel Aviv University Faculty of Law and is a member of Israel’s National Committee for Genetic Experiments in Human Beings. David Chinitz, Ph.D., received a PhD in public policy analysis from the University of Pennsylvania and is Senior Lecturer in Health Policy and Management at the Hebrew University-Hadassah School of Public Health as well as scientific chair of the European Health Management Association.
years have supported the right of payers to ration health care. This throws the issue back to the legislature for an expression of social preferences regarding rationing. Given that the legislature is loathe to be seen as rationing health care, the search is on for a silver bullet solution, and so enters quality. However, it is doubtful whether quality basics can be defined well enough for the issue of health care rationing to be kept out of the courts in the future. Elsewhere it has been argued that technical tools for rationing will always have to be complemented by an institutional framework that mediates the political conflicts that inhere to health resource allocation decisions.

Israel provides a case that illustrates these issues quite explicitly. The 1994 National Health Insurance Law (NHI) includes, among its central provisions, the establishment of a highly specific and detailed basket of health services to which every citizen is entitled, as well as a process for updating this basket. At the same time NHI requires the government to provide the funding necessary to implement the basket and any updates. As both the contents of the basket and the amount of available funds are subject to political processes, this setup constitutes a visible, if not necessarily self-conscious, attempt to blend technocratic and political inputs within a binding legal framework.

This arrangement has been overlaid on a society that has traditionally been characterized by high respect for the legal system, an increasingly activist judiciary, and high levels of litigiousness together with a cultural affinity to medicine and a liking for new technologies. Israel, well known for being among the countries with the highest ratio of physicians to population, has an equal ratio for lawyers. Indeed, the NHI was proposed and given impetus by a State Commission of Inquiry headed by a Supreme Court Justice. One can note here a built-in tension: a desire to create a system that summons minimal judicial intervention in a society that leans towards such intervention.

This paper describes and analyzes the evolution of this setup over time, focusing especially on the role of the courts. Given the above outline of the NHI setup, one might hypothesize that the courts would play a limited role relative to the executive and legislative branches. In general, the description and analysis below support this proposition. At the same time, the article suggests that the political-bureaucratic framework of NHI is vulnerable enough to shocks in either individual demand for health services or government fiscal policy that consequent judicial intervention may have a strong influence on the system.

Section 1 provides additional background on the Israeli system necessary for understanding the role of the courts. Section 2 reviews significant cases and assesses the impact of court rulings on health policy. The discussion section summarizes the Israeli case while the conclusion draws lessons for other countries.

1. The Israeli Health Policy Context
Like many health systems in the developed world, the Israeli system has been undergoing a clearly identified and significant structural reform during the last decade. This reform has been overlaid on evolutionary changes in Israeli society and in the health system whose roots go even deeper. An example is the increasing perception of the citizen as a consumer and of patients as clients, along with growing competitive behavior on the part of health providers and insurers. Moreover, the health system reforms coincided with a constitutional revolution which articulated certain basic human rights and recognized the power of judicial review over acts of legislation.

The Israeli health care reform has been described and analyzed elsewhere, and for the purposes of this paper only those aspects most relevant to the role of the courts will be discussed. As mentioned above, perhaps the most central provision of the 1995 NHI Law is the guarantee of a defined basket of services to all citizens based on medical need. In addition, citizens are entitled to join any one of four health funds and to switch funds once a year without limitation. Beyond a health tax paid to the National Insurance Institute, citizens are not required to pay any fees for coverage of the basket. Health funds receive age adjusted capitation payments from an earmarked health tax, and are forbidden to demand additional payments for the basic coverage. Health funds contract with providers, though government intervenes in determining health plan payments to general hospitals as described in more detail below.

What is important to emphasize here is that this setup blends a competitive approach with significant government and legislative intervention. In effect the only competition allowed over the basic basket is linked to perceived quality of service. Health funds may add services to the basic basket, but their incentive to do so is tempered by the fact that, aside from not being able to compete on price, adding services may incur adverse selection. Health funds are permitted to provide sup-
plemental policies that cover items not included in the basic basket. These provisions serve to highlight the question of what is included in the basic basket, and the mechanism by which the latter is updated.

The NHI Law sets out a statutory mechanism for the update, involving the Ministers of Finance and Health, as well as an advisory Health Council made up of multi-sectoral representatives. Each year government allocates a defined amount of money for additions to the basket. A public committee made up of representatives of various health system stakeholders and public interest representatives recommends a prioritized list to the Minister of Health. Needless to say, the available allocations always fall short of the constant flow of new technologies competing for entry into the basket. The process is based on significant technical input (health technology assessment) but is also highly political. In March 2005, the Prime Minister personally intervened to add 150 million NIS (about $30 million US) in order to expand the number of drugs and treatments added to the basket.

As in many cases, the Israeli reform can be looked at as a social learning process regarding the difficult decisions concerning allocation of health resources. The question is whether the process is indeed one of actual positive learning, or an entropic one in which social consensus evaporates and the public health system is undermined. In this connection the courts play a key role, either contributing to the solidity of the system or serving as the arena for resolving tough conflicts that perhaps would be better solved politically. In the following section we describe the role of the courts in order to assess the contribution they make to the system.

2. Role of the Courts
In addition to defining substantive rights to health insurance and care, the NHI Law defined several rights of a procedural nature, such as the right to receive information from health funds about services and service providers, and the right to submit a complaint to a national ombudsperson appointed under the Law (hereinafter – the NHI Ombudsperson). It also guaranteed the right of insured individuals to bring an action in a court of law and vested the labor courts with jurisdiction over all disputes between insured individuals and the health funds (s. 54).

Jurisdiction
The venue of the labor courts was chosen because they also have jurisdiction under social security law, which defines the test of residence for inclusion in the NHI scheme. In addition, labor courts are considered to be a forum that is more friendly to individuals than the civil courts and accessible even without representation by lawyers. However, in the decade since the enactment of the Law there has been relatively little litigation in the labor courts, probably no more than a few hundred cases, compared to thousands of actions each year in labor relations and social security claims, and compared also to over 3,000 complaints submitted each year to the office of the NHI Ombudsperson.

One possible explanation, from the perspective of litigating lawyers, is that in most cases the value of the individual claim would not justify the costs of litigation, and there is no provision in the Law for class actions. But another factor to be considered is that people do not know they have a right to seek relief in court. A national survey on NHI rights literacy conducted in 2002-3 found that 57% of the respondents knew they had a right to appeal health plan decisions, but only 1% knew they had a statutory right to take action in the labor court, compared to 18% who knew they could submit a complaint to the NHI ombudsperson.

The jurisdiction of the labor courts is limited to disputes between insured individuals and the health funds, or providers of services included in the basic basket, acting on their behalf. So far litigation in the labor courts has, for the most part, addressed refusals of health funds to provide services to insured individuals. The key question that came before the courts is whether the claimed services are or are not included in the basic basket, or whether they are included for the plaintiff’s specific clinical indication. The jurisdiction of the labor courts does not, however, extend to challenges to the actual content of the basic basket. In other words, the labor court will adjudicate an individual claim of entitlement to a particular treatment, but cannot question the composition and scope of the basic basket of services. This matter would come before the Supreme Court within its jurisdiction to exercise judicial review over the executive branch of government.

Public Funding
The power to update the basic basket is ministerial, and the decision whether or not to expand the basic basket and to add new technologies is an executive decision. As such, it is subject to judicial review by the Supreme Court sitting as High Court of Justice. Until recently, this issue had not been brought to court. But the Supreme Court had occasion to address several disputes between the health funds and the treasury, regarding the calculus of public finances and funding, in which it is generally reluctant to interfere.

Two cases with large financial consequences addressed global agreements for the provision of hospital services to insured persons. The NHI Law had subjected the health funds to strict administrative and fis-
eral hospitals owned by the health fund. Under the legislative intention, and refused to intervene. It is not surprising that the Court is hesitant to intervene in such technical and obscure fiscal matters. The legislation is a rubber stamp on numerical equations, which reflect the results of a certain bargaining process, with trade-offs against other legally unrelated chips. These are mega-deals in terms of Israel's health economy. The Court is not necessarily well equipped to deal with accounting. And it has a long standing position of refusing to grant relief and intervene in executive decisions concerning matters of resource allocation. In short, the Supreme Court does not interfere with the state budget. This is left to politics.

The same reticence is evident in a case in which one of the health funds challenged a foundation of the budgetary framework of the NHI – known as the "cost of health index." Section 9(b)(1) of the NHI Law provides that the "cost of the basket of health services shall be updated annually by the rate of increase in the cost of health index." The Fifth Appendix to the Law defined the index as a computation of several other indices, including health sector salaries, consumer prices, wholesale pharmaceutical prices, and construction costs. But the Law also provides, in section 9(c), that the "Minister of Health and the Minister of Finance may, on the recommendation of the Health Council and with the approval of parliament's Labor, Welfare and Health Committee, change the composition of the cost of health index, or update the cost of the basket of health services due to demographic changes in the population." The Health Council had indeed discussed the matter and recommended that the index be changed to account for an increase in the size of the population and the aging of the population, and also to account for the high costs of new medical technologies and drugs. The Minister of Finance, however, refused to exercise his authority to change the index, claiming budgetary constraints and statutory discretion.

In Maccabi Health Services v. Minister of Finance, the health fund challenged the minister's discretion in refusing to change the index, arguing that it would not be able to fulfill its statutory duties to provide the basic basket of services within a balanced budget. While the case was pending in court, a parliamentary committee of inquiry produced a report which made a similar rec-
ommendation to that of the Health Council – to enact an automatic increment to the “cost of the basket,” rather than leave the decision about the specific amount in the hands of the government, as part of the general negotiation around the annual budget that takes place at the end of each fiscal year. The cabinet discussed this parliamentary recommendation, but decided to continue the present policy, that is, to leave the question of the amount allocated each year for updating the basic basket with new technologies to the political process, within the framework of budgetary priorities.

Against this background the Supreme Court dismissed the minister’s argument about budgetary constraints. It expounded a theory of certain statutory duties that must be carried out by the state and cannot be dependent on budgets, it being the responsibility of the authority holder to find the fiscal resources. When it comes to substantive duties, the Court said, the state cannot refuse to allocate resources on grounds of priorities. *Ubi jus remedium ibi* – where there is a right there is a remedy, and not the other way around. The more specific the right, in terms of nature and scope, the less room for arguing budgetary constraints. Thus, the court said, the government could not evade its duty to allocate resources to the basic basket of services which all residents have a right to receive on the basis of medical need. Therefore the government must either find an appropriate budgetary framework, or amend the Law and change the duties it lays down. The NHI Law, said the Court, laid down a structure of rights and duties, and was intended to replace the previous system of subsidies in which the state would, from time to time, without any regularity, bail out health funds in deficit with retrospective hand-outs in exchange for agreement to recovery plans.

As for the finance minister’s discretion, the Court discussed at some length the statutory status of the Health Council as an advisory body, which at the very least obliges the ministers to give due consideration to its recommendations, and to make a reasoned decision after having done so. However, after all this lofty talk, the Court decided not to order the minister to act upon the Council’s recommendation, commenting, “[w]e have never instructed the state to pay anyone out of its budget amounts of such magnitude as those that the health funds are asking for in the present matter.”

Three Basket of Services

As mentioned above, the general jurisdiction over rights of access to health care lies in the labor courts. From the outset it was clear that the basic basket of health services set a minimum standard of legal duty. In one of the first cases relating to this matter, *Medzini v. Clalit Health Fund*, the plaintiffs, suffering from multiple sclerosis, challenged an internal administrative restriction on the number of physiotherapy treatments included in the basic basket. The court upheld the restriction, but dismissed the argument of the health fund that it may not fund treatments that are not included in the basic basket. The court stated that the services which the health fund is obliged by statute to provide are a “floor,” and that it may provide additional services. As opposed to social security rights, the labor court ruled that the health funds have the authority to provide and finance health care beyond the call of duty, as a matter of *cheded* – grace, kindness, compassion, mercy, or equity.

At the same time, the registered clinical indications for drugs are substantive, that is, there is no duty to provide and finance services for indications that are not registered. This was upheld by the Supreme Court in its role as High Court of Justice. The courts do not, accordingly, usurp the statutory mechanism or authority for extending the scope of the basic basket; that is, they do not undermine the ministerial, political process. Despite this, on occasion the labor courts have ordered health funds to provide uncovered treatments as a temporary measure of interim relief, making it difficult to limit treatment subsequently. Also, as we shall see, the labor courts have developed a jurisprudence of procedural justice for answering the needs of patients for health care in exceptional cases.

The NHI Law established a clear mechanism for amending the basic basket. According to statute, nothing can be added to the basket without guarantee by the Ministers of Health and Finance that adequate budget resources are available. Nothing can be removed without approval of the parliamentary committee on health and social welfare. The Health Council is to serve as an advisor to the Minister of Health, among other things, on the contents of the basic basket. As a result of processes described elsewhere, a public committee was established in 1998 under the auspices of the Health Council, to deliberate and make recommendations regarding additions to the basket (hereinafter – the Basket Committee). The committee operates in view of a defined amount of money made available by the Ministry of Finance for this purpose. Of course the demands for additions to the basket exceed the available funds.

In the years following the enactment of NHI, only one new drug was added to the basket, subject to clinical guidelines that were more restrictive than the registered indications for use. Patients suing for the drug obtained access to it outside of the guidelines issued by the MOH, since these guidelines had, according to the court, no legal status. As a result, clinical indications which serve as a rationing mechanism for public funding of new technologies, are now incorporated in the
ministerial orders that are promulgated as secondary legislation to update the basic basket. This implies that medical advice can only serve as a basis for a claim to a service if that advice conforms to the list of treatments specified in the Law including the guidelines and indications for use. As such, appeals to court intervention to overturn health fund decisions to deny access are unlikely to succeed unless the letter of the law has been ignored. The courts acknowledge budgetary restraints and accept standards of evidence-based medicine as benchmarks for public funding.

Recently, however, increasing media attention and public debate have resulted in challenges to the political decision making process about budgets for adding new services and drugs. During the 2005 budget cycle, the Basket Committee demanded additional funds so as to avoid the need to prioritize what it considered to be life saving technologies, and the public pressure resulted in added expansion of the basic basket. Subsequently, too, certain individuals denied access to new expensive drugs which were not added by the public committee filed a petition with the Supreme Court – still pending – which challenges the actual discretion of the Basket Committee.

The results of the Supreme Court deliberations on these suits will be of particular interest from a theoretical and practical point of view. Action research conducted both in the style of public deliberation by one of the authors as well as surveys of citizen attitudes conducted by the other reveal surprisingly high levels of understanding and support for the current process. Over time the public increasingly expresses a preference for quality of life enhancing treatments relative to treatments aimed at extending life in the case of incurable diseases, potentially providing a cushion for difficult decisions faced by the public committee (See Table 1). Citizens also express relatively high levels of trust in health institutions as arbiters of health policy relative to other institutions (See Table 2).

A judicial determination that the current process is not acceptable could endanger the learning process evolving around definition of the basic basket. At the same time, it is not clear that the political and bureaucratic factors surrounding the process are contributing to preserving and building the existing levels of trust. This is a classic example of a moment when the balance of influence may shift among the political, bureaucratic, and judicial branches, an issue that will be taken up in the discussion section.

**Equity in Exceptional Cases**

It is inevitable that cases will arise in which the definitions of services included in the basic basket will not suffice to determine whether a given service should be provided to a particular individual in need. As specific as clinical guidelines may be, they are always based on aggregated data and there will always be individuals who defy the statistics. The courts seem to have sought to walk a fine line between not interfering with administrative decisions and responding to the sometimes difficult to ignore, requests of patients for intervention. Sometimes the courts prod, but do not insist, that the health plans go “beyond the call of duty” in meeting the needs of patients outside the strict letter of the law.

Regional courts, which are the courts of first instance, are generally more responsive to patient claims, but precedent is set by the more distanced appellate instance – the National Labor Court. For example, in several instances regional courts granted interim relief for patients seeking life extending drugs to which they were not strictly entitled, and once they received them it was difficult to reverse the situation and cease the treatment regime. Consequently the National Labor Court set down rules for granting interim relief, which take into account the severity of the medical condition and the danger of deterioration if not treated, as well as the urgency of the need for treatment. Most importantly, however, the National Labor Court has developed rules of procedural fairness for internal health fund decision making in exceptional cases, which must be exhausted before patients seek judicial intervention.

Thus, in *Elad Shitrit v. Meuhedet Health Services*, a patient with a rare condition sought funded access to hair removal treatment. The claim was denied in the regional labor court but upheld in the National Labor Court on grounds that the health fund’s decision making process had been flawed. Standards of due process,

---

**Table 1**

<table>
<thead>
<tr>
<th>Domain of Health Service</th>
<th>Year of Survey</th>
<th>1997</th>
<th>2001</th>
<th>2003</th>
<th>2004</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transplantation</td>
<td></td>
<td>8.8</td>
<td>7.9</td>
<td>7.8</td>
<td></td>
</tr>
<tr>
<td>Expensive Treatments</td>
<td></td>
<td>8.25</td>
<td>7.6</td>
<td>7.7</td>
<td>7.5</td>
</tr>
<tr>
<td>Terminal Conditions</td>
<td></td>
<td>7.9</td>
<td>6.8</td>
<td>6.5</td>
<td>6.7</td>
</tr>
<tr>
<td>Quality of Life (e.g., physiotherapy)</td>
<td></td>
<td>7.8</td>
<td>7.7</td>
<td>7.7</td>
<td>7.9</td>
</tr>
<tr>
<td>Nursing Care</td>
<td></td>
<td>8.1</td>
<td>7.4</td>
<td>7.5</td>
<td></td>
</tr>
<tr>
<td>Fertility Treatments</td>
<td></td>
<td>7.7</td>
<td>7.5</td>
<td>7.2</td>
<td></td>
</tr>
</tbody>
</table>

**Explanation:** Four random, representative samples of the Israeli public over 18 years of age were surveyed in the years indicated and asked to rank groups of health vignettes on a scale of 1-10. The average scores, indicated in the cells, reduced over time, especially regarding preference for treatment of terminal conditions.
drawn from maxims of natural justice that are well established in administrative law, were explicated to include: the duty of the health fund to specify the reasons for its decision; the duty to investigate the facts and merit of each case; the duty to keep records of any discussion of the case, to disclose all relevant information to the patient, and to grant a hearing; and the opportunity to make arguments and present one’s case before health fund decision making bodies.

As a result, all of the health funds have established special mechanisms for addressing claims made by members in exceptional circumstances, and at least one has even formulated substantive criteria for making decisions in such cases. In Maccabi Health Services v. Ada Ben Zvi, the National Labor Court reiterated that health funds provide health services as public authorities acting by virtue of statute, and hence must abide by the standards of public administrative law which mandate reasonableness, fairness, equity, and human dignity. But it then held by a majority ruling, that it would not intervene in the professional judgment of an internal health fund appeals committee which had reached a decision not to fund a life-extending cancer drug. The considerations weighed by the health fund were as follows: Is the treatment life saving? Is there evidence in the literature for the efficacy of the treatment in question, or is it experimental? Is the treatment approved in other Western countries? Are there alternative effective treatments that are included in the basic basket of services? What are the overall priorities of the health fund? And what are the budgetary implications for the health fund given that the health fund might have to make treatment available in other similar cases?

Indeed, the provision of a service on the basis of exceptional need raises a quandary of equity: a decision to provide a service to one patient, on grounds of compassion rather than statutory right, might open the door for other similarly situated individuals to invoke a right to the same service, claiming non-discrimination. Hence, the health funds complained for a while that answering the needs of patients in exceptional cases would result in an expansion of the basic basket, since they are under a statutory duty to refrain from discriminating amongst their own members. This argument was however settled in the case of Clalit Health Fund v. Gil Ayal Dekel. There the plaintiff claimed a right to have the health fund provide a pulmonary device for home use, while it was the health fund’s policy to purchase the device for public clinics. The Tel Aviv regional labor court found that one or two members of the health fund had received it for their personal use as exceptions to the policy, and admitted the plaintiff’s claim on grounds of discrimination. The National Labor Court, however, overturned this decision, explaining that the norm of equality applies to the health fund even in relation to services that are not included in the basic basket, but that the present case ought not be compared to the previous exceptions but to the policy itself – which was not flawed. The fact that the health fund had made an exception to its policy in a singular case did not mean that it had discriminated against the present plaintiff. And one could not infer a new implicit policy which differed from the previous express one, from the fact that the device was provided in isolated cases. In other words, exceptional decisions do not trump overall policy.

To summarize, the courts acknowledge the need to balance cost and clinical utility in setting priorities within budgetary constraints, and seem to accept standards of evidence-based medicine as indicators of reasonableness in making funding decisions. But in the most heart-wrenching cases, reason and compassion may be at odds. For one final example, consider the
case of Meirav Binyamini v. Clalit Health Services, which concerned a 35 year old woman, the mother of two small children, who was terminally ill with breast cancer, and sought funding from the health fund for an experimental treatment which involved among other things bone marrow transplantation, and would cost initially about $50,000. Section 3(d) of the NHI Law provides that insured individuals have a right to health services of “reasonable” quality. The Haifa regional labor court interpreted the standard of reasonableness as setting a scientific evidence-based standard of care, and with great reluctance reached the conclusion that the efficacy of the requested care was as yet unproven. Since the treatment in question was experimental, there was no medical justification to force the defendant to fund the treatment. Nonetheless, poignantly, the court summed up its decision thus:

In view of all that has been said, with a heavy heart and with no other choice, it is our duty to act and judge according to the law and not according to the inclinations of our hearts and our sympathy with the plaintiff. There is no legal possibility to oblige the defendant to undertake the burden of financing allogenic bone marrow transplantation, because this treatment is not included in the basket of health services according to the Law. Notwithstanding, we find it appropriate to appeal to the defendant to examine with an open heart and a willing soul, and to see whether there is any possibility to assist the plaintiff with full or partial funding of the treatment that is requisite for the allogenic bone marrow transplantation, in view of the fact that this is the only treatment left that might save her life!

3. Discussion
Two main issues have surfaced in the evolution of the courts’ role in the Israeli health system. First, do the courts increase access to health care; and, second, is the process balanced in terms of the relative influence of the courts as opposed to the legislative and executive branches?

Taking the second issue first, the thrust of the above review of cases is that the courts are clearly inclined to resist overturning administrative decisions of health system institutions. In the two main arenas where the courts could intervene, namely, macro budgetary policy and disposition of individual appeals regarding denial of specific services, the courts have shied away from taking an activist role. Even in cases where the court sees fit to express explicit dismay at the suffering caused by resource allocation decisions, only clear violation of existing statutes will induce judicial relief. It appears that the Israeli system has constructed a tight and clear set of laws and procedures regarding the health system, leaving options for major change largely in the hands of the political system.

This should not necessarily be taken as failure of the courts to increase access to health services. Providing succor to individual plaintiffs may not always be perceived as increasing overall social access to health care. What is granted to one may come at the expense of others. The Israeli courts seem to have integrated trade off thinking into deliberations regarding the health system. There is recognition that budgetary constraints are a relevant and admissible aspect of public decision making. In addition the courts recognize the necessity and right of the government and delegated entities such as the health funds to set priorities, implying that some services that would be possible to provide with potential benefit to someone will not be allocated resources.

The main concern or intent of the courts, even if not at the outset of the NHI regime, appears to be to ensure that decision making processes meet standards of fairness and rationality in the sense of considering pertinent points such as the availability of substitutes for the medical intervention under deliberation.

There are those who suggest, as one option, more “judicial legislation” in the sense of having the courts interpret the intent of health legislation in terms of certain values. It is by no means clear that this is a desirable direction for the health system and for Israeli society overall. The court should not be called upon to fill in policy gaps that should more appropriately be dealt with through political processes supplemented by technical input such as health technology assessment. We also draw attention to the evidence presented above indicating that the Israeli public, while placing relatively high trust in the judicial system, places even more trust in the MOH and the health funds to handle health policy decisions.

4. Conclusions
Among society’s unavoidable “tragic choices,” allocation of scarce health care resources may present the most difficult challenges. Until perhaps the beginning of the 1990s, these choices were largely made implicitly in most developed health systems, at either the macro budgetary or micro “local justice” levels. However, the confluence of increased cost consciousness with advances in health services research, in particular the pathbreaking work of John Wennberg on small area variations, has led a number of countries into the waters of explicit rationing and priority setting.

Notwithstanding the move toward more explicit decision making in the allocation of health resources, it appears that in few, if any instances, are stakeholders,
especially politicians, eager to be held accountable. Thus in many countries there has been observed a cycling among an array of decision tools such as devolved responsibility,\textsuperscript{30} evidence based decision-making, clinical guidelines, global budgeting, the legal system, etc. These tools are attractive because they provide the appearance of being technocratic and politically neutral.\textsuperscript{31} Some might say that this is the best society can do in this tough policy realm.

Be that as it may, the Israeli case, for better or worse, appears to have pushed the notion of explicit rationing quite far, mainly by pitting budgetary limits against priority setting in the context of a universal health insurance regime. The role of the courts is perhaps the best litmus test for this. The cases reported and analyzed above suggest that the legislation regulating the health system is “tight” enough that there is relatively little room for judicial discretion. Decisions regarding both the size of the budget and its allocation among competing needs are squarely located in the legislative and executive branches, constituting a classic case of combining politics and technocracy.

This is not to say that the system is perfect. While evidence shows that the Israeli public is gaining awareness of the process and sensitivity to the dilemmas involved, more could be done to keep the public informed. The make-up of the Basket Committee and turnover among its membership could be altered. Disclosure of the access of various interest groups, such as pharmaceutical companies, to the process and their influence could be increased.

In this context, the courts play an important role. Even when not overturning administrative decisions, consideration of various claims gives them more visibility. The courts’ assertions regarding the clear, if sometimes sad, realistic limitations imposed by the NHI arrangement, serve as a source of education regarding the health system. This is a valuable contribution to the viability of the system and potentially to increasing trust in the system.

Do the courts, by playing this role, advance access to health care? While some would quickly argue the negative, we suggest that the procedural focus of the courts makes an important contribution to the successful management of the health system. This, in turn, hopefully, can be a basis for increasing public trust in and social solidarity regarding the health care system. Whether this will occur depends on the bureaucratic competence of government and its faithfulness to underlying social values,\textsuperscript{32} and, given persistent media and anecdotal evidence regarding significant bureaucratic hurdles facing patients seeking access to covered services, is certainly worthy of further examination. Increased trust and solidarity, in and of themselves, probably contribute to better health outcomes, if not directly to better access to health care. They also contribute potentially to better access by smoothing the difficult allocation decisions confronting society in the realm of health. But this outcome is not guaranteed and the courts may find themselves recalled to deal with problems of access to health care.

Notwithstanding the nearly unavoidable difficulties of dealing with problems of access to health care, for a country confronting numerous “tragic choices,” health policy has turned out to be a possible model. Citizens of Israel might hope that lessons from health care might inform policy making in other key areas. For other countries, the Israeli case provides one, perhaps paradigmatic model, of striking a balance among technocratic, political, and judicial inputs to managing access to health care.

References


6. Indeed, some observers quip that each Israeli physician can be sued by exactly one Israeli lawyer.


10. HJC Meuchedet Health Fund v. Minister of Finance (Judgment of September 9, 2003, as yet unpublished).

11. HJC 2725/02 Chalit Health Services v. State of Israel (Judgment
of September 8, 2004, as yet unpublished).

12. HCJ 2344/98 Maccabi Health Services v. Minister of Finance 54 (5) P.D. 729.

13. DBA 5-7/97 Yirma Medzini v. Clalit Health Fund, 33 PDA 193.


17. C. Shalev and D. Chinitz, supra note 16.


22. AA 1091/00 Elad Shitrit v. Meuhedet Health Fund 35 PDA 5.


24. AA 1267/01 Clalit Health Fund v. Ayal Dekel 37 PDA 97.


32. R. G. Evans, supra note 5.
Accountability for Rationing – Theory into Practice

Christopher Newdick

Most now recognize the inevitability of rationing in modern health care systems. The elastic nature of the concept of “health need,” our natural human sympathy for those in distress, the increased range of conditions for which treatment is available, the “greying” of the population; all expand demand for care in ways that exceed the supply of resources to provide it. UK governments, however, have found this truth difficult to present and have not encouraged open and candid public debate about choices in health care. Indeed, successive governments have presented the opposite view, that “if you are ill or injured there will be a national health service there to help; and access to it will be based on need and need alone.” And they have been rightly criticized for misleading the public and then blaming clinical and managerial staff in the National Health Service (NHS) when expectations have been disappointed.

Although UK “politics” obscures the issue, a significant contribution to the hard choices debate has been made by health policy analysts. Two currents of debate have developed. First, substantive, mechanisms for allocating resources have been suggested, most notable amongst which is the utilitarian concept of the Quality Adjusted Life Year (QALY) which maximises the return from finite resources by favouring treatments on the basis of the quality of life gained, the length of time in which it will be enjoyed, and its cost. The other side of the debate emphasises procedural solutions to the problem, based on due process and accountability. This argues that that there is no right, or wrong answer and that our responses and instincts will change over time. The preferable solution is not based on substantive benefits based on a mathematical equation, but in the way decisions are made. Recent interest has been shown in refining the ingredients of “process” and the manner in which it promotes reasonableness, accountability and public confidence.

If the debate was ever polarized, we have now surely found a middle way where both sides recognise the other’s contribution; for example that whilst QALYs cannot dictate outcomes, they are certainly a relevant factor in the decision-making process. Central to the debate are questions such as who makes decisions, according to what processes and criteria, and do patients and the public play a role? The following considers responses to these questions in the NHS and the difficul-

Christopher Newdick, J.D., is a barrister and the Reader in Health Law at the University of Reading, UK. He has served on the Department of Health’s Medicines Commission, is an honorary consultant to Reading PCT and member of the Berkshire Priorities Committee. His book Who Should We Treat? – Rights, Rationing and Resources in the NHS was published in 2005.
ties and limitations that have been encountered. We discuss: (1) judicial responses to patients dissatisfied with resources allocation decisions and (2) implementation of policy in the NHS by examining (a) a local initiative in the county of Berkshire, UK, as one example of the values of fair and consistent resource allocation, and (b) two further national initiatives which guide and shape the process. We will see that as a general rule the UK prefers procedural solutions, but there has been a significant exception with respect to the creation of the National Institute for Health and Clinical Excellence.

Judicial Responses
In England and Wales, legal challenges to resource allocation are dealt with by judicial review, in which the courts scrutinise the propriety of public authority decision-making and, if needs be, refer the matter back to the authority to be reconsidered in the light of the court’s observations. The effectiveness of the remedy lies in the willingness of the courts to criticise those who have been appointed, often by virtue of their special expertise, to make these difficult decisions. Until the mid-1990s, judicial review in England provided an inadequate remedy to review health care resource allocation because it was too passive and deferential. In a most egregious case decided in 1988, the health authority refused to fund the drug beta interferon for the treatment of multiple sclerosis (MS).11 Given its profile of clinical benefits and financial costs, and exercising the discretion it believed was available to it, the health authority refused to fund the drug. The applicant suffered MS and challenged the decision.

In England and Wales, legal challenges to resource allocation are dealt with by judicial review, in which the courts scrutinise the propriety of public authority decision-making and, if needs be, refer the matter back to the authority to be reconsidered in the light of the court’s observations.

In judicial review, the court confirmed that the Secretary of State for Health issued a health service circular recommending that health authorities consider using the drug beta interferon for the treatment of multiple sclerosis (MS).11 Given its profile of clinical benefits and financial costs, and exercising the discretion it believed was available to it, the health authority refused to fund the drug. The applicant suffered MS and challenged the decision.

In judicial review, the court confirmed that the Secretary of State’s advice imposed no obligation upon the health authority to fund the drug. In principle, it is for each PCT to decide whether, and how to implement national policy. Nevertheless, the case succeeded on another ground. The health authority had failed properly to bear in mind a relevant consideration in coming to its decision, namely the NHS policy set out in the circular, in the absence of which its decision was irrational. It was not obliged to adhere to national policy, but neither was it entitled to ignore it. This emphasises the duty to take into account all the relevant considerations, including governmental guidance. For example, under “Choose and Book,”12 the Labour government has introduced a policy which permits patients a choice of hospital in which to be treated, and the power to choose a date on which they will be admitted for treatment. This policy is not mandatory on the health service, but the government’s aspiration must be taken into account properly by local decision-makers.

In legal theory, this detaches the Secretary of State for Health from responsibility for the NHS. The practical reality, however, is rather different. Central control of the NHS by the Department of Health is often enforced by (non-legal) “targets” imposed centrally. “Perfor-
manoeuvre indicators” are crucial to the way in which local managers are assessed and managed by the centre. Those that perform poorly under intense pressure to achieve the objectives set. Those that fail may be forced to resign from their posts. Extra-legal they may be, therefore, but the impact of central policy on NHS management is difficult to under-estimate. Of course, the disadvantage of a system which gives so much informal leverage to the centre is that short-term political objectives may undermine the longer term policies of local decision-makers responding to needs of their communities. For example, in vitro fertilisation, is not a high priority in many PCTs. Yet, in 2004, the Secretary of State endorsed guidelines recommending that such services should be provided. Cynics will suspect that the pressure to respond to this particular group, rather than, say, geriatric patients, or those with long-term mental illness, is driven by electoral anxieties, rather than responsible public health policy. It may pressurise PCTs to allocate resources in ways that do not reflect local priorities and work against local accountability for decision-making.

2. The courts have explained some of the criteria of reasonableness and emphasised that decision-makers must pay particular attention to the impact of decisions on individual patients. In R v. NW Lancashire HA, ex A. D. & G., three applicants for transsexual surgery were refused funding for their care. They appealed against the decision on the ground that the merits of their case had not been properly considered. The Court of Appeal upheld their case and quashed the decision. The court acknowledged that hard decisions between competing demands are unavoidable and that a policy to place transsexualism low in an order of priorities of illnesses for treatment and to deny it treatment save exceptional circumstances...is not in principle irrational, provided that the policy genuinely recognises the possibility of there being an overriding clinical need.

The process of doing so should consider relevant circumstances, such as the nature and seriousness of each type of illness, the effectiveness of various forms of treatment for it, and its cost. But such a process should always recognise the individual patient and be sufficiently flexible to accommodate those with exceptional need. In the circumstances of this case, the applicants had been subject to a blanket ban on the procedure which led the health authority to ignore the possibility of their having exceptional merits. For this reason the decision was over-turned and returned to the health authority to be taken again.

Each PCT should consider how this principle should be incorporated locally. Equally, with so many PCTs there is potential for significant differences in policy between them. In a national health service is this desirable? The preferable solution is for PCTs to enter consortia as a means of understanding and refining the differing solutions that exist to the challenge of resource allocation. Whilst each PCT would retain responsibility in law for its own decision-making processes, a more refined approach would emerge, well-informed by the evidence of good practice elsewhere. Ironically, the obstacle to such a development may be government. Whilst the Department of Health at the centre of the NHS refuses to concede the inevitability of hard choices in health care, local PCTs will not be encouraged to admit the need for procedures to manage them. Despite this, some local initiatives of this nature have been undertaken and we discuss below the collaborative response of the PCTs in Berkshire, Oxfordshire and Buckinghamshire.

3. What about treatments where evidence of efficacy is equivocal because systematic and reliable clinical data is unavailable? The matter was considered in the transsexuals’ case. Recall that the authority refused the patients access to surgery. One of its reasons for doing so was that there was insufficient evidence of clinical efficacy to support the procedure. Exhaustive randomised trials were not available because none had been conducted. How should health authorities assess a new technology in the absence of authoritative trials? The court held that it is not reasonable to exclude all such treatment from consideration because it is unrealistic to require cogent clinical evidence in every case. Time may not have permitted completion of proper trials. In this case, despite the absence of reliable evidence, a body of doctors regularly performed transsexual surgery and supported making it available to suitable patients. Where a body of opinion supports the use of a technology, “it is...not open to a rational health authority simply to determine that the procedure has no proven clinical benefit while giving no indication of why it considers that is so.”
Thus, even in the absence of reliable clinical evidence and notwithstanding the differences of view that existed, the authority was duty-bound to give the proposed treatment proper consideration and to assess it according to the fair and consistent values discussed above. A pressing example of the difficulties associated with this duty arises in connection with “orphan drugs.” Orphan drugs are intended for the diagnosis, prevention or treatment of rare life-threatening or chronically debilitating conditions affecting not more than five in 10,000 persons. Given the numbers of patients concerned, they may generate small revenues for manufacturers and an abbreviated, less expensive licensing procedure has been introduced to encourage their development. Obviously it is difficult to accumulate long-term clinical-studies of such drugs. Inevitably, however, once the manufacturer’s licence is granted, the problem of clinical uncertainty arises. Because the demand for the drug is small, the costs of manufacture may be high. And because the licensing process is abbreviated, the quantity and quality of clinical evidence may be limited. Should the treatment be provided?

It seems natural to say that those with rare medical conditions for which a “promising” intervention exists should have access to it. On the other hand, PCTs will say that each such intervention has huge “opportunity costs.” Should PCTs in effect take responsibility for the long-term clinical assessment of such drugs? And from whom should resources be diverted to accommodate the costs? When hard choices have to be made, many PCTs will be reluctant to divert resources to expensive treatments of uncertain clinical value. This leaves a conundrum; orphan drugs may not be funded because of lack of clinical evidence, but there is lack of evidence because the treatment is not funded. Nevertheless, under the transsexuals’ case, there is no duty to fund such treatment provided the principles governing reasonable procedures are followed.

4. Finally, say a patient has been refused access to care under a general priorities framework. Health authorities have a duty to consider whether the patient merits exceptional access to care. The procedure involved in “exceptional cases” was considered in R (on the application of F) v. Oxfordshire Mental Healthcare NHS Trust, in which the Oxfordshire Priorities Forum refused a patient access to the “ideal” psychiatric treatment of her choice which was vastly more expensive than the “reasonable” treatment proposed by the health authority. This “second-best” decision was challenged by way of judicial review. The court emphasized the need for balance between fairness and openness as to the reasons for the decision, and the danger of turning the hearing into a trial. Fairness required that the claimant should be able to tell the Forum in writing why she believed that resources should be allocated to her. However, the process need not take place in person, or be in the nature of a judicial hearing. The court said that the meeting of the forum was essentially a discussion between medical experts, not a contested hearing. “Decisions on funding affect lives, not just liberty. That is not a good reason to judicialize them. They are agonizingly difficult decisions, and they will not be made any easier or better if they are encumbered with legalistic procedures.”

This recognizes the breadth of discretion available and the importance of fair procedures. But it also provokes the question: how extensive is the duty to provide “exceptional” appeals procedures? The process carries its own resource implications. Clinicians adjudicating over exceptional cases appeals are not treating patients. Were the process to expand significantly, and larger numbers of patients were to resort it, objection would be raised that doctors were spending disproportionate time hearing appeals about patients rather than treating them. This has its own “opportunity costs” and suggests that the appeals process itself may have to recognize the need for priority setting.

Since 1997, therefore, the courts have significantly refined the expectations of PCT decision-makers and, by doing so, improved patients’ procedural rights. Equally, they have recognized the legitimacy of rationing and tend not to express preferences for some patients, or treatments, over others. Substantive merits have been left firmly in the hands of the NHS.

Implementation of Policy

Let us now turn to the manner in which interest in accountability for reasonableness has been implemented within the NHS. We examine (a) the local response of the PCTs in the county of Berkshire by their creation of a Priorities Committee working within an Ethical Framework and (b) two national responses: (i) the creation of the National Institute for Health and Clinical Excellence, and (ii) the commitment to enhance public involvement in the NHS.

(a) A Local Response: The Berkshire Priorities Committee

The county of Berkshire contains six PCTs (each with population of about 100,000 people) which have created a “Priorities Committee” to advise them on matters of resource allocation. The BPC has no independent statutory authority and its purpose is advisory only. Nevertheless, it has created a consistent basis on which resource allocation decisions may be made and its advice is generally adhered to. The committee is guided by its Ethical Framework which evolves over time and is set out below.
Berkshire Priorities Committee, Ethical Framework

The Berkshire Priorities Committee (BPC) is a committee of representatives of NHS organizations, including Community Health Councils in Berkshire. Its purpose is to advise the NHS in Berkshire as to the interventions and policies that should be given high or low priority. PCTs are under a statutory duty to promote the health of the local community. They are also under a duty not to exceed their annual financial allocation. This inevitably means that, from time to time, hard choices have to be made. Recommendations from the Berkshire Priorities Committee are intended to help PCTs choose how to allocate their resources to promote the health of the local community.

The BPC has developed this Ethical Framework to enable it to make fair and consistent decisions which treat patients equally. Many of the decisions it makes will involve the exercise of judgment and discretion and there will be room for disagreement both within and outside of the BPC. Although there is no objective, or infallible measure by which such decisions can be based, the Ethical Framework enables decisions to be made within a consistent setting which respects the needs of individuals and the community.

The BPC believes that people have equal rights of access to health care. There may also be times when some categories of care are given priority in order to address health inequalities in the community. However, the Priorities Committee will not discriminate on grounds of personal characteristics, such as age, gender, sexual orientation, race, religion, lifestyle, social position, family or financial status, intelligence or cognitive functioning.

The BPC will assess a patient’s health needs according to their capacity to benefit from health care. In the absence of evidence of health need, treatment will not generally be given solely because a patient requests it. Similarly, a treatment of very little benefit will not be provided simply because it is the only treatment available. This is necessary to ensure that resources are used to provide the greatest health benefit.

The Ethical Framework is especially concerned with the following:

1. Evidence of Clinical Effectiveness
   The Priorities Committee will seek to obtain the best evidence of clinical effectiveness. It will promote treatments for which there is good evidence of clinical effectiveness. It will not normally recommend treatment that is shown to be ineffective. When assessing evidence of clinical effectiveness the outcome measures which will be given greatest importance are those considered important to patients. Reliable evidence will often be available from large-scale randomized clinical trials. Evidence may also be available from less authoritative sources and this will also be considered. Patients’ evidence of significant clinical benefit is also relevant.

2. Cost of Treatment
   Because each PCT is duty-bound not to exceed its budget, the cost of treatment must be considered. The cost of treatment is significant because investing in one area of health care inevitably diverts resources from other uses. A single episode of treatment may be very expensive, or the cost of treating a whole community may be high. The Priorities Committee will compare the costs of treatment to its overall benefit, both to the individual and the community. It will consider technical cost-benefit calculations, but these will not by themselves be decisive.

3. The Need For Health Care
   The Priorities Committee will consider the health needs of patients according to their capacity to benefit from health care. So far as possible, it will respect the rights of patients to choose between different treatment options, subject to the support of the clinical evidence. Urgent and life-saving treatment will be given a high priority, as will treatment which effectively treats “life time,” or chronic conditions such as arthritis, mental illness, or sensory impairment. When evidence of clinical effectiveness is equivocal, options for treatment will be given particular attention.
   There will be no blanket bans on treatment since there may be cases in which a particular patient has special circumstances which present an exceptional need for treatment. Each case of this sort will be considered on its own merits in the light of all the clinical evidence. PCTs may create procedures to consider such exceptional cases on their merits.

4. Needs of the Community
   Public health is an important concern of the BPC and it will always seek to make decisions which promote the health of the entire community. Some of these decisions are promoted by the Department of Health (such as the guidance from the National Institute for Health and Clinical Excellence (NICE) and National Service Frameworks). Others are produced locally. The BPC also supports effective policies to promote preventive medicine which help stop people becoming ill in the first place.
   Sometimes the needs of the community may conflict with the needs of individuals. Decisions are difficult when expensive treatment produces very little clinical benefit. For example, it may do little to improve the patient’s condition, or to stop, or slow the progression of disease. Where it has been decided that a treatment has a low priority and cannot generally be supported, a patient’s doctor may seek to persuade the PCT that there are exceptional circumstances which mean that the patient should receive the treatment within procedures established by the PCT.

5. National Standards
   The Department of Health issues guidance and directions to NHS bodies which may give priority to some categories of patient, or types of treatment. These may affect the way in which health service resources are allocated. The BPC operates with these factors in mind and recognises that its discretion may be affected by National Service Frameworks, NICE technology appraisal guidance, and NHS directions.

This framework seeks to combine the approach recommended by the courts with good ethical practice. It promotes consistency but the generalized nature of the principles cannot disguise the fact that decision-making is not driven by analytical or algebraic precision. Clinical evidence is central to the process, but it provides only a platform on which differences of opinion often exist and discretion is exercised. The Committee is composed of around 20 people. It is chaired by a chairman of one of the PCTs and attended by doctors from primary and secondary care, PCT directors of finance, senior hospital managers, those responsible for purchasing (or “commissioning”) patient care, representatives of patients interests, an ethicist and a lawyer. Estimates of QALY costs are very important to the committee, but they are not decisive. The committee meets monthly.
Beneath the Ethical Framework hang over 80 separate policies on specific treatments. These provide advice and guidance to clinicians consistent, where relevant, with national policy. Treatment may be recommended to particular categories of patients, or in limited circumstances only. A policy may recommend that alternative treatment is attempted before use of more expensive therapy. Each policy states either that the treatment is recommended for use within the NHS, or that it is a low priority. In the latter case, additional funding will not normally be provided to accommodate it. Policies are regularly reviewed in the light of new clinical evidence, or national policy. The Priorities Committee manages a small notional budget of funds set-aside by its six PCTs in order to ease the cost of drugs during the financial year in which funding decisions are made. Thereafter, the continuing costs of the treatment are met from within the allocations of each of the PCTs.

Consistent with the idea that patients should not be excluded from treatment by blanket bans, the framework incorporates exceptional cases review committees (ECRC). Patients who have been refused access to care as a result of a general policy introduced under the Ethical Framework may argue that their specific circumstances merit an exceptional response. Exceptional cases often concern matters connected with cosmetic surgery (for example in connection with breast reduction/augmentation and pinnaplasty for children’s ears). Adjudication over exceptional cases is the responsibility of each PCT on the recommendation of ECRC staffed by clinicians not previously involved with the case under appeal.

This system was introduced in 1999. It has the advantage of transparency to patients and encourages consistency across the county. Indeed, the system is being extended to Oxfordshire (which has had a similar ethical framework since 1999) and Buckinghamshire as the “Thames Valley Ethical Framework.” Patients may not be pleased that hard choices on health care resources are unavoidable, but they may be reassured by a system which seeks to ensure that the process is conducted openly, fairly and consistently between them. Note, however, that even this three-county system operates in only 16 PCTs out of a total of 303 in England. Clearly, if it is a success, similar frameworks and collaborations should be encouraged throughout the NHS.

(b) Two National Responses
There have been two further developments relevant to priority setting at national level: (i) the National Institute for Health and Clinical Excellence (NICE) and (ii) the movement to enhance public involvement with the NHS. We consider each in turn.

(i) The National Institute for Health and Clinical Excellence:
NICE was created to reduce the inevitable post-code differentials in access that can arise when so many health authorities have statutory independence. Its statutory duty is to advise PCTs whether to fund a particular technology taking account of “the promotion of clinical excellence and of the effective use of available resources in the health service.” As we have seen, as a general rule, the Secretary of State may encourage and persuade PCTs to respond to problems in a certain way but he has no authority to insist that they do so. The exception is that he may issue “Secretary of State’s Directions.” Until the creation of NICE, directions had seldom been used in matters of NHS resource allocation. Secretary of State’s directions now give NICE Technology Appraisal Guidance (NICE TAGs) the status of directions to PCTs. The directions require that

...a [PCT] shall, unless directed otherwise by the Secretary of State...apply such sums amounts of the sums paid to it...as may be required to ensure that a health intervention that is recommended by [NICE] in a Technology Appraisal Guidance is, from a date not later than three months from the date of the Technology Appraisal Guidance, normally available [to patients].

In this way, NICE may mandate that resources are set aside by PCTs to fund the costs of its guidance. NICE, therefore, is not a procedural solution to the resource allocation question; it provides legal rights to treatment and it offers the opportunity to consider a number of problems associated with substantive guarantees to health care.

Although NICE considers the resources available to the NHS, the manner in which it does so remains unclear. It says that decisions on cost effectiveness must include judgments on the implications for other healthcare programmes that may be displaced by the adoption of a new technology. On the other hand, it continues: “[NICE] does not consider the affordability of the new technology but does account of how its advice may enable the more efficient use of healthcare resources.” What does this mean? NICE guidance is based on the cross section of clinical evidence, from large-scale, randomised trials, to the anecdotal experience of carers. As the cost of the treatment increases, high quality evidence becomes the more important. The World Health Organisation (WHO) has urged NICE to state whether it uses a cost threshold above which approval will not normally be given. NICE responds, however, that it does not use an upper financial limit. Although it becomes cautious before recommending treatments
with QALY costs above £30,000,\textsuperscript{29} persuasive evidence of significant benefit may justify approval. As cost increases so “the case for supporting the technology...has to be increasingly strong.”\textsuperscript{30} Indeed, in the case of an effective treatment for myeloid leukaemia, NICE has approved treatment (imatinib) with a cost probably in excess of £60,000/QALY.\textsuperscript{31}

Inescapably, the funding implications for PCTs are considerable. Since the cost of NICE TAGs has to be provided within existing budgets, PCTs have to disinvest from some areas in order to accommodate others. This “blanket” duty to pay for NICE guidance is questionable. NICE makes recommendations as to clinical effectiveness. But evidence of efficacy is entirely different to the question whether it should take automatic priority over everything else. This is the impact of the mandatory status of NICE guidance on PCTs. Yet the government provides no assistance on the manner in which disinvestment should take place. Many have been unable to divert sufficient resources to provide uniform adoption of NICE guidance. For example, wide variations appear to exist in the uptake of NICE guidance on the treatment of cancer.\textsuperscript{32} Notwithstanding the very difficult situation in which PCTs are placed, this failure is in breach of the mandatory duty imposed by the Secretary of State’s Directions and renders them amenable to judicial review. NICE is beginning to articulate the social and ethical values that drive its decision-making which might assist these highly sensitive decisions. However, like Berkshire’s Ethical Framework, this is a very difficult enterprise and the result is unsurprisingly imprecise.\textsuperscript{33}

Note too that although NICE TAGs have mandatory impact on PCT “commissioners,” they remain wholly discretionary for doctors. NICE explains the need to retain clinical discretion at the doctor-patient level on the basis that although doctors should take its guidance fully into account, “it does not override their responsibility for making appropriate decisions in the circumstances of the individual patient...because even the best clinical guideline is unlikely to be able to accommodate more than around 80% of patients for whom it has been developed.”\textsuperscript{34} The clinical response to NICE guidance, therefore, should be very different to that of the PCT. This means that, given variations in clinical practice between doctors, there will always be unevenness in the adoption of NICE guidance. Some doctors may be unsure about the guidance itself, or consider their patient to come within the exceptional minority.

Perhaps NICE is a double-edged sword. It certainly encourages greater national harmony in the uptake of NHS treatment. But its guarantees of access for some impose greater strain on the availability of treatments that have not received NICE TAG approval.

(ii) Public Involvement
We now turn to the issue of public involvement in the NHS. This clearly seeks to refine decision-making processes, rather than guarantee access to specific treatments. The statutory duty on PCTs and NHS hospitals to involve the public in decision-making is:

To make arrangements...that persons to whom...services are provided are directly or through representatives involved and consulted on (a) the planning of the provision of those services, (b) the development and consideration of proposals for changes in the way those services are provided, and (c) decisions to be made by that body affecting the operation of those services.\textsuperscript{35}

Public consultation is a generic goal not limited to the issue of priority setting. Nevertheless, it is especially appropriate to the sensitive questions provoked by health care resource allocation. But it also poses a number of difficult problems. As the Department of Health concedes: “there is no consensus about what ‘public involvement’ means...When different people collaborate in the pursuit of public involvement, they may or may not have a shared understanding of the concept.”\textsuperscript{36} For example, who is “the public?” Patient pressure groups exist to promote the interests of those they represent. They provide invaluable insights into the nature of illness, the suffering it causes and the treatments available. They are an essential part the understanding needed by PCTs. But they are often single-issue groups that do not claim to represent the public. Were they to be given disproportionate influence, resources might be diverted from other deserving, but less articulate groups. Much care must be taken balancing the views of influential groups with those who are unlikely to express themselves clearly or at all.\textsuperscript{37}

The Court of Appeal has insisted that for the process to be genuine, consultation must be undertaken before the final decision has been made, i.e. at a time when proposals are still at the planning stage; it must include sufficient reasons for particular proposals to allow people to properly consider and respond to it; adequate time must be given for this purpose; and the product of consultation must be conscientiously taken into account when the ultimate decision is taken.\textsuperscript{38} A failure in this regard will render the process amenable to judicial review. But this still leaves very broad discretion to individual PCTs and hospitals as to when, how and with whom consultation should be conducted.

And what should be the persuasive value of such a consultation? Say a PCT wishes to close an old and well-loved hospital. Standards of safety may be impossible to maintain given the limited numbers and cross-
section of patients admitted for care, or the site may be more valuable if sold in order to generate additional funds for a new, more efficient hospital, more appropriately located for patients. Or the closure is a central part of a strategy for the entire region. The need to do so may be clear at the “macro” level, yet local people may resent and oppose it. Despite their opposition, the legal authority for decision-making rests with the board of each PCT, not with its advisory committees, or “the public.” PCT boards are duty-bound to bear in mind such local views, but fairness, consistency, and long-term planning (as well as the law) depend on their commitment to the community as a whole. In these difficult cases, therefore, local people may never be satisfied that their views and objections have been given proper weight.

Various mechanisms have been attempted for consulting the public, including health forums, citizens’ (or patient/carer) panels, focus groups and newsletters. Each carries its own, sometimes significant costs. NICE has introduced the first national body of this type; its “Citizen’s Council” has 30 members specifically created to consider “what should NICE take into account when making decisions about clinical need?” The Council does not advise on the clinical and scientific merit of matters brought before NICE. Its function is to consider “the hopes, values and priorities of people generally.” Such a body demonstrates NICE’s commitment to public involvement, but what impact should a body of non-elected individuals have on national policy? It recommends, for example, that “the response to clinical need should be to provide without prejudice or political, geographical, economic, or social preference, the best care or treatment that can be provided within the resources available, to make the most improvement to a patients’ well-being” and that “factors such as the individual’s value system and their cultural and religious views should be taken into account.” These sentiments would surely attract general support but they provide limited guidance on many of the difficult choices dictated by scarce resources. This suggests that the NHS still has much to learn both as to the manner in which these difficult choices should be managed. Matters of this nature are inevitably “political” and government should play a role (along with local people) in dealing with them. Many public health systems acknowledge this and UK governments should do the same. Nevertheless, until we have NHS-wide guidance, PCTs should devise local mechanisms for doing so; in particular they should form consortia as a means of learning from one another, promoting best practice and increasing equity and consistency throughout the NHS. Given the difficulties associated with substantive solutions to the problem, I have suggested that procedural responses based on an ethical framework are most likely to provide the way forward.

Critics of the procedural response to health care priority setting may say that it is so “polycentric” and imprecise as to be meaningless. No matter how much information is made available, choices will tend to be based more on intuition than fact and logic. Much better, therefore, to abandon the pursuit analytical purity and, instead, to “muddle though elegantly” doing as best we can to manage the swamp of clinical, ethical, social and policy options that present themselves. However, as we come to trust government less and demand greater transparency from decision-makers, this is unlikely to be persuasive. “Muddle” could permit invisibility as to the reasons for decisions, inconsistency between patients and irrationality in decision-making. Guided by the courts, the Frameworks being developed in Berkshire and NICE recognise the need for consistency, a balance between individual rights and the needs of the community and that even if decisions are to some extent intuitive, we can explain why they have been made and demonstrate that like cases are treated alike.

As we have seen, many of the solutions, both procedural and substantive, generate problems of their own so the process will evolve over time. Public engagement is unlikely to provide easy answers because it too presents as many problems as it solves. Nevertheless, although health authorities may respond in different ways to the problems they face, they should be guided by a common commitment to fair, open and consistent values and procedures, capable of withstanding the scrutiny they will rightly attract.

**UK governments may be reluctant to muddy their hands in the health care rationing debate, but PCTs have no such liberty.**
References
2. See the criticism by Professor Sir Ian Kennedy in Learning from Bristol: The Report of the Public Inquiry into Children’s Heart Surgery at the Bristol Royal Infirmary, 1984-95 (London: Cm. 5207, 2001): 57, at paragraph 31.
4. However, QALYs may under-value other ethical objectives, such as equality, clinical need and ability to benefit. See N. Daniels, “Four Unsolved Rationing Problems: A Challenge,” Hastings Centre Report 24, no. 4 (1994): 27.
8. See National Health Service Act 1977, s 97, as amended.
9. S 1, National Health Service Act 1977: The duty is imposed on the Secretary of State and delegated to PCTs.
15. Ibid, at 408.
19. Ibid, paragraphs 77 and 80.
20. Exceptionally, in Simms [2003] 1 All England Reports 669, the court encouraged use of experimental treatment for patients suffering CJD.
25. Secretary of State’s Directions of 2003. NICE guidance remains discretionary in Wales.
41. Ibid, at 14.
42. Ibid, at 19 and 22.
Access to care has become a key and contentious issue in the Canadian health care system. In this article, I explore the role of Canadian courts in determining rights to access public health insurance (Medicare), beginning with a brief overview of the Canadian system and its distinguishing features, and then moving to discuss challenges to governmental limits on publicly-funded Medicare using the Canadian Charter of Rights and Freedoms. I argue that the Canadian courts are not, as is often charged, proactive in this area. I question whether the deference exhibited by courts to governmental limits on Medicare is justified given concerns about the fairness of the principles and processes followed by decision-makers. In sharp relief to the judiciary’s conservative approach to applications for better or timely access to publicly-funded Medicare is the recent Supreme Court of Canada’s decision in Chaoulli v. Quebec (Attorney General) which upheld a right to buy private health insurance for “medically necessary” hospital and physician services. This decision overturned legislation in the province of Quebec prohibiting private health insurance. This case raises the prospect of future Charter challenges to provincial legislation designed to prevent the flourishing of a private sector. It seems that Canadian courts are destined to play a significant role in determining the future balance between public and private insurance in the Canadian health care system. Sadly, however, Canadian courts to date have staked out only a very limited role to help improve the internal fairness and operation of publicly-funded Medicare itself.

The Canadian System
Canadians have historically been both protective and proud of their publicly-funded health care system, known as Medicare. Popular sentiment maintains that “universal publicly funded health care is part of what it means to be a Canadian and reflects our core values,” and Roy Romanow in his 2001 report on the future of Medicare said Canadians believe access to medically necessary health care services is a “right of citizenship.” Indeed the joke often goes that Canadians are just like Americans; except for the reality of good health care. In contradistinction to the United States’ Medicare program, Canadian Medicare is a universal program, funded through general taxation revenues, covering all its citizens and not just those over 65 for “medically necessary” hospital and physician services. In addition, there is much more private financing for services falling outside the “core” of hospital and physician services.

Colleen M. Flood, J.D., is the Canada Research Chair in Health Law and Policy, Faculty of Law, University of Toronto.
example, in 2004 over 52% of total spending on drugs came from the private sector.⁵

The Canada Health Act,⁶ the main federal legislation in this area, protects by requiring that all “medically necessary” hospital services and “medically required” physician services be fully publicly-funded. Canadians, proud of their Medicare system, often assume that rights to health care are enshrined in the Canada Health Act. In fact, the Act is a mere spending statute by which the Federal government, through transfer payments, attempts to entice the 10 provinces to comply with certain minimum requirements in their respective health insurance plans.⁷ If the provinces don’t comply with these requirements, then the federal government may withhold transfer funds.⁸

The most important and controversial feature of the Canada Health Act, and what distinguishes Canada from most other jurisdictions, is the requirement for first-dollar public coverage of all hospital and physician services defined as “medically necessary.” This is in contrast to countries like the UK, Australia, and New Zealand, where those with the means to do so are free to purchase privately physician and hospital services. In other words, Canadians with means have historically not been able to buy their way to the front of queue for medically necessary hospital and physician services. The goal of Medicare has been that access to care be allocated on the basis of need, and not ability to pay. Across the 10 provinces in Canada, private markets in hospital and physician services are effectively stymied through prohibitions on private insurance and/or prohibitions on the ability of physicians to work simultaneously in both the public and private sectors.⁹ Systems in the UK, New Zealand, and Australia, that allow people to purchase private insurance to buy better or faster treatment for services that are ostensibly covered by the public system, provide what is known in Canada as a “two-tier” system.

Until the recent decision of the Supreme Court in Chaoulli, such a system has been an anathema in Canadian policy circles because of the deep and enduring support there is on the part of Canadians for Medicare. The health care system, premised on the core values of social justice and solidarity, is in a sense an extension of the ideals that underlie the fabric of Canadian society. Medicare promotes equality and the community rather than the rights of individuals, and to Canadians represents one of the major distinguishing features between themselves and Americans. The Chaoulli decision has, however, opened up the debate about the role of private insurance and the private sector in hospital and physician services. To be clear, Canada does have and has always had a significant private sector (over 30% of total funding for the system comes from the private sector) but private spending is centred on prescription drugs and home care and not on essential hospital and physician services.¹¹

Health care systems around the world, whatever their public/private configuration, are facing enormous fiscal challenges. Rising expectations, expanding technologies, and the increasing costs of labour (training and remuneration of health professions) are raising sustainability questions in all systems. Canada is no exception, and that nation’s commitment to what the Europeans call solidarity is increasingly under strain. In particular, the Canadian system is being passively privatized as the emphasis in health care moves from physician and hospital services to newer technologies. Prescription drugs needed outside hospital walls, medical equipment, and genetic therapies all fall beyond the protections of the Canada Health Act and into a world of mixed public and private financing. But it is these new technologies, like drugs and genetic therapies, which are the future of health care and are the fastest growing component of health care expenditures.¹² The net result is a system where a drug like insulin is not consistently publicly-funded but, for example, annual general check-ups, for which there is no evidence of any health benefits accruing, are fully publicly funded because they are classified as “medically necessary” physician services.¹³

Thus the pressing challenge for Canadians in terms of access and sustainability is how to broaden the range of services and treatments covered in the publicly funded system, and narrow it at the same time. In other words, in theory at least, Canadians should not indiscriminately fund all physician services and neither should they publicly-fund all drugs – but rather only those which are truly medically necessary. Canadians should, ideally, be able to treat like medical needs alike regardless of whether the optimal therapy is a hospital service, a drug, or a genetic therapy. But of course this begs the question of how they decide and who decides what services are medically necessary?¹⁴ And when resources are limited and cannot fund all “medically necessary” services or needs, how do they prioritize? Beyond this, should people be able to purchase privately services of a higher quality or services more quickly

The Canadian system is being passively privatized as the emphasis in health care moves from physician and hospital services to newer technologies.
than what is offered in the public system and, if so, what will be the consequences for the public system?  

**Charter Challenges**  
In the face of governmental limits on public funding of care, growing expectations of treatment, increasing technologies, and limits on the ability to purchase the same health services privately, Canadians are increasingly turning to the courts to resolve their access concerns. The most high-profile cases have revolved around challenges using the *Canadian Charter of Rights and Freedoms*.  

Since 1982 the *Charter* has protected Canadians’ rights and freedoms by limiting the ability of governments to pass laws or take actions that discriminate or infringe on human rights. Section 52 provides that the constitution of Canada is the supreme law of Canada, and that any law inconsistent with it is of no force or effect.  

With respect to possible challenges by citizens to governmental limits on publicly-funded health care there are three relevant sections: 7, 15 and 1. I will explore each section in turn.

**Section 7**  
Section 7 provides “Everyone has the right to life, liberty and security of the person and the right not to be deprived thereof except in accordance with the principles of fundamental justice.” Generally, section 7 has (at least to date) been interpreted in a way so that governments do not have a duty to provide publicly-funded health care. There is, however, one exception involving a lower-court decision in 2004 concerning the rights of women to access privately-funded abortion services in a timely fashion. In *Doe et al. v. The Government of Manitoba* a court found that the complainants’ section 7 rights had been breached as a result of the government’s failure to fund abortion services outside of public hospitals (for example at a private clinic) thus forcing the women, who could not purchase services privately, to queue in waiting lists in a public hospital.

The *Chaoulli* decision, however, has yet to be tested on appeal. Even if it stands, it could be distinguished as applying to the special situation of women needing an abortion. In the case of abortion the consequences of delay or waiting even a month are patently clear. Demonstrating the physical risks and psychological harms of waiting a month, for example for a hip operation, may be more difficult.

But this interpretation of *Doe v. Manitoba* and indeed the future of Canadian Medicare itself has now been put in question as a result of the Supreme Court’s decision of *Chaoulli* in June 2005.
end result. Progressives are dismayed to see the Charter being used to justify privatization of Canada’s major program of redistribution (from the rich to the poor; from the healthy to the sick). Regardless of which side of the fence one sits on vis-à-vis the pros and cons of private vs. public health insurance, many have criticized the majority’s high-handed treatment of complex policy questions. Through their comparative analysis of health care systems, the majority amply demonstrates why courts should be extremely cautious of wading into these difficult policy choices.

The fundamental error that the majority makes is to conflate all health care systems with some role for private insurance into one group; namely two-tier or parallel-private systems. But in fact there are at least four distinct ways of financing health care that incorporate both public and private sources: Parallel public and private systems (e.g., the UK); Co-payment (e.g., New Zealand); Group-based (e.g., Germany); and Sectoral (e.g., Canada). In particular, the majority unfavorably compares Canada to European countries such as the Netherlands and Germany. In Canada, certain health care sectors are entirely publicly financed (hospital and physician services), whereas other sectors (like prescription drugs) are financed privately either through insurance or out-of-pocket payments. In group-based systems, which operate in the Netherlands and Germany, most of the population is covered through social insurance but certain higher income groups are left to purchase private insurance to cover all their needs (and this in turn is often highly regulated to ameliorate access and equity concerns). In parallel public and private systems (with allows complementary or duplicate insurance), all are covered by the public system but the purchase of private insurance is permissible and can lead to faster or higher quality care in the private sector than that which is available in the public sector.

To be clear, a European kind of system, steeped in history and with years of experience in regulating the private sector to achieve social justice and efficiency goals, will not evolve in Quebec as a result of the Supreme Court’s usurpation of the legislative ban on private health insurance. The private system that could emerge in Quebec after Chaoulli will also be primarily for higher-income individuals, but that is where the similarity with European group-based systems will end. Rather than creating an opportunity for more people to access timely treatment, the parallel private system that emerges will succeed only in taking away limited resources from the public system and transferring them to a small number of people. Several studies provide evidence that an unregulated parallel private health care sector will not improve and may lengthen waiting time for those who are left in the public system.

The Chaoulli decision means great uncertainty for the future of Canadian Medicare. Strictly speaking, the Chaoulli decision is limited to Quebec but it has implications that flow far beyond those borders. Privateers are marshaling their resources across the country to challenge similar laws in other provinces. Every province is anticipating Charter challenges to the myriad of laws presently in place that cumulatively suppress a flourishing private sector in health care. It is unclear how the court would respond to these other legislative provisions that effectively dampen the ability of a private sector to flourish; e.g., provisions preventing physicians working in both the public and the private sectors simultaneously or regulations that prevent a physician charging a private fee that is higher than the public tariff for the same service. As litigators for those who want more freedom to provide private health care and aggrieved patients marshal their legal resources, provinces across the country are considering their options. Some are seeking guidance on how to better insulate themselves from review; others, such as Alberta, have welcomed the Chaoulli decision as a way to attack the universality principle of the Canada Health Act.

Section 15 and Section 1
The Chaoulli decision speaks to rights to purchase private health insurance. Returning to the issue of access to public treatments, most of the Charter litigation in the context of contesting entitlements to publicly-funded care has occurred in the context of section 15. Section 15 (1) provides that “Every individual is equal before and under the law and has the right to the equal protection and equal benefit of the law without discrimination and, in particular, without discrimination based on race, national or ethnic origin, color, religion, sex, age or mental or physical disability.”

Even if a law is found to be discriminatory under section 15 (1) it may be “saved” by section 1 of the Charter. Section 1 provides that the Canadian Charter of Rights and Freedoms guarantees the rights and freedoms set out in it subject only to such reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society.” Essentially section 1 allows a government to defend a policy or decision found to be discriminatory by demonstrating that, although considered, the needs of those discriminated against were out-weighed by other pressing needs or considerations.

There have been a number of section 15 challenges that have grabbed the public’s attention, and there has certainly been an increase in the number of applicants before the court making these sorts of challenges. But between 1985 and 2002, of the 33 cases that have challenged health care policy, only 11 have been successful. Most of the successful cases have been about doctors
contesting governmental limits on their ability to set up practice wherever and whenever they like.27 Of the cases that have challenged policies limiting insured medical services only one, Eldridge v. British Columbia,28 has been upheld at the Supreme Court level. As discussed earlier, Doe v. Manitoba,29 addressed these issues at the lower court level, and may yet be appealed.

The Eldridge case,30 decided by the Supreme Court in 1997, was a claim by a deaf couple that they were discriminated against in contravention of their section 15 rights by the failure of the British Columbian government (and more specifically a hospital) to fund interpretation services. The facts of this case were very compelling.

Linda Warren, a deaf woman, was giving birth to twins and there were problems at the time of delivery. A nurse communicated to her through gestures that the heart rate of one of the babies had gone down. The twins were whisked away from the distressed woman who did not know what was going on apart from a note being flashed at her with the word “fine” written thereon. The claimants argued that failure to provide an interpreter impairs the ability of deaf patients “to communicate with their doctors and other health care providers, and thus increases the risk of misdiagnosis and ineffective treatment.”31 In effect, hearing impaired persons receive lesser quality health care services. Hearing impaired persons have an equal right to the benefits of Medicare, but when the government (or an agent providing government objectives such as a hospital) fails to pay for an interpreter, these rights are inadvertently discriminated against. The Supreme Court found that the claimants were indeed discriminated against and directed the British Columbia government to provide interpretation services in a manner consistent with the equal protection benefits espoused under section 15 (1).

There were two ways to interpret the Eldridge decision. The first interpretation is that it should be limited to the extent that the applicants were not seeking any new health treatments per se but the right to communicate and receive information about the same health treatments that those without hearing loss enjoyed. The Supreme Court strongly emphasized that effective communication is an indispensable component of the delivery of a medical service.32 This interpretation would significantly narrow the precedent of this decision applicable to future Charter challenges. The second interpretation, and one promoted by some in the media, is that the Eldridge decision represented unprecedented judicial activism and threatened to open the floodgates to the use of the Charter by those unhappy with government limits on public financing. But this judicial activism critique, insofar as it applies to Charter challenges to governmental limits on Medicare, is wrong.33 Rather, the courts have adopted a more conservative approach in this area, as the following examples demonstrate.

A case that more clearly illustrates the conservative judicial approach taken to these types of claims is the 1990 case of Brown v. British Columbia (Minister of Health).34 In that case, the drug AZT was placed on the Pharmacare plan which meant that all AIDS patients apart from those on social assistance or residing in long-term facilities had to pay for part of the drug’s costs. The British Columbia Supreme Court found there was no violation of the Charter as the reason that AIDS drugs were on a different plan from cancer drugs had to do with the differences between the drugs and not inequality as contemplated by section 15.

Another example of the court’s conservative approach to access claims using the Charter is the 1999 case of Cameron v. Nova Scotia Government.35 This was a claim for public funding for a particular form of in vitro fertility treatment, intra cytoplasmic sperm injection (ICSI), which is the standard treatment option in cases where the male partner is infertile. On appeal, the majority of the Nova Scotia Court of Appeal found the government’s funding decision contravened the claimants’ equality rights under section 15(1), but the violation was justified under section 1. The majority held that IVF and ICSI are medically necessary services for infertile individuals, the denial of which offended their dignity,36 but that such discriminatory action was nonetheless saved by section 1 of the Charter, being a reasonable limit prescribed by law in a free and democratic society.

In its section 1 reasoning, the court stressed the historical and on-going financial pressures faced by the Nova Scotia public health system, and concluded that the objective of controlling health care spending was pressing and substantial. The majority emphasized the fact that ICSI had not yet gone through the necessary processes to determine its safety and efficacy, that al-
though ICSI was not funded there were other services, financed for the infertile, and that funds not spent on ICSI were spent on other pressing health care needs. The court thus concluded that the government had satisfactorily discharged the onus upon it to demonstrate that the discrimination was justified. As Kent Roach has observed, “the Court will often uphold legislation if the government does a half-way decent job of mounting a section 1 defence.”

It is a salutary warning to advocacy groups considering Charter challenges that the Cameron decision seems to have been used as a justification by Canadian governments not to revisit the issue of funding IVF treatments. Over the course of the last decade, many jurisdictions including Australia, Israel, and New Zealand have moved to fund publicly a range of IVF treatment, including ICSI, in some circumstances. The adverse finding in Cameron seems to have set back infertility advocates in their claims for public support much more than may have been the case if the Charter challenge had never been brought.

Much of most recent complaints about judicial activism in Medicare have focused on the trial and appeal decisions in the case of Auton v. British Columbia. This case involved a claim on the part of parents with autistic children for public funding of a relatively controversial treatment, Lovas therapy, which is claimed to have dramatic benefits for autistic children. At trial, the court overturned the British Columbia provincial government’s decision not to fund the treatment, and this was upheld on appeal. But this decision was overturned at the Supreme Court in November 2004.

In its reasoning the Supreme Court took a formalistic approach. It focused in specifically on whether the claimed benefit was a benefit provided by law as required by section 15. Normally the requirement for a claimant to demonstrate unequal treatment under the law is not a key issue as the claim relates to a clear statutory benefit or burden. The Supreme Court found that the lower courts mistakenly concluded that the legislative scheme requires the provision of all medically required treatments. What the provincial legislation does do, as required by the Canada Health Act, is to require the provision of all medically required treatments provided in hospitals and by physicians. The province, at its discretion, then funds some non-core services by naming classes of specific health care practitioners in regulation, like dentists, optometrists, etc. The therapists who provide Lovas therapy are not listed. The court then says to think otherwise would be to mean that all non-core services needed by those who fall within the enumerated or analogous grounds would have to be covered.

With respect, although the outcome of the decision may be fair, the court’s reasoning is disappointing. The applicants had to surmount the insurmountable in being required to establish discrimination in the processes by which some classes of services provided by providers, like optometrists are listed and not others. The court required the applicants to provide evidence of discrimination that was distinct from the mere fact of exclusion from a partial social program. The claimants needed evidence that other treatments with similar doubts about efficacy were being funded for other groups. But this begs the question of how such evidence is to be obtained when decision-making processes are deeply embedded within governmental bureaucracies and there is no transparency as to the criteria guiding decision-making?

Another concern with the court’s approach is the short shrift given to the issue of the severity of underlying medical need. The Supreme Court claimed that the right comparator group was “emergent new technologies,” in other words, those technologies that had yet to be proven effective through randomized controlled trials. In sharp distinction to the lower courts – where the fact that the treatment may help autistic children avoid a life-time of exclusion and suffering weighed heavily – the Supreme Court ignored the issue of the severity of medical need in question. When considering what principles should guide decision-making, obviously lack of evidence of efficacy must be considered but another relevant consideration must surely be severity of medical need.

Is Deference Required?

The Supreme Court’s decision thus starkly raises the extent to which courts should defer to the black-box of government decision-making. There is a danger that
decisions are made on the basis of what serves the short-term political interests of governments and major stakeholders, such as the medical and nursing associations. The court in Auton had a golden opportunity to send a strong signal to decision-makers that they must articulate clear and fair principles when determining what services attract public funding. The unfairness of the system is underscored by the fact that there are many hospital and physician services and treatments that are fully publicly funded and yet there is no evidence of their efficacy; for example, annual general check-ups. Many services that are fully publicly funded are in pursuit of outcomes that are far less important than curing autism. In order not to be discriminatory, and to provide equal treatment as required by section 15, like needs should be treated alike. Formalistic divisions between categories of services and providers should be no defense to discriminatory treatment.

In the wake of Auton, one can see that the claim that the courts are opening the floodgates of access to Medicare through Charter challenges is manifestly not true. The courts are extremely reluctant to wade into the fray and in particular to create new entitlements. Even when a breach of section 15 is found, little is needed for a court to be persuaded that the deprivation of the right is justified under section 1. But is such deference on the part of the courts towards governmental decision-making warranted?

In a three-year program of research funded by the Canadian Health Services Research Foundation, my colleagues and I have explored how decisions are made in Canada about what is publicly-funded. Our research raises concerns about the principles that presently drive decision-making. For example, our exploration of the Byzantine process for determining what physician services are publicly funded in the province of Ontario reveals decisions are a result of negotiations between the provincial government and the medical association which acts as the bargaining agent for physicians in Ontario. The negotiations are over the level of tariffs for each medical service. Thus the schedule of tariffs becomes the list of services that are publicly funded and defined as “medically necessary.” So the concept of “medical necessity” does not drive the determination of what is or is not publicly funded. Rather it is a label that is applied ex post labor negotiations.

One would hope that decisions about the boundaries of publicly-funded Medicare are some function of values (including Charter values), available resources, severity of medical need, and evidence about the relative costs and benefits of treatments. It seems, however, that choices in Canada are more a function of accidents of history and long-held accommodations between governments and the medical profession, inflexible regulations and law, and the result of turf protection and lobbying by different stakeholders and interest groups.

It is clear that the principles that presently guide decision-making about the limits of Medicare are unfair; but it is, of course, much more difficult to determine the principles that should guide decision-making and to rank these principles or values should they conflict. For example, many may argue that a key principle guiding decision-making should be evidence of efficacy. Indeed, applying the best evidence would seem a sine qua non for the best possible decisions. But caution is required, for such an approach may mask certain biases; for example, preferring randomized controlled trials at the expense of other kinds of evidence (such as that acquired by Aboriginal peoples in the use of traditional medicines). It may also discount the relevance of severity of medical need and the need for “curing” kinds of services that have no measurable effect on health outcomes (e.g. palliative care). A demand for evidence may mean that treatments for rare diseases, for example Fabry’s disease, may never be approved because of the difficulty of running randomized controlled trials across extremely small populations. Legal academics may tend to promote a focus on the welfare of the most vulnerable with a particular emphasis on Charter values. Public health experts may prefer to focus on the health of the population as a whole; economists will take a similar population perspective focusing on measurable health outcomes. By contrast clinicians may argue for not only future benefits but severity of medical need. All of these principles and values are reasonable, but ultimately, when in conflict, which of these should have priority?

International experience, from the Netherlands, Oregon, New Zealand, and Israel, suggests that the project of determining an explicit list of principles to guide decision-making is often doomed to failure because of the intractability of ranking those principles. Balancing competing and conflicting values would not seem such an insurmountable challenge from a legal perspective – jurists must often do so. But political institutions – charged with making politically-sensitive trade-offs between people and services – have either proved unwilling or unable.

More recently, in the academic literature there has been a shift of focus to the processes of decision-making. The rationale is that fair processes can legitimate the decision. This perspective, first expounded by Norman Daniels and James Sabin, has become known as “accountability for reasonableness.” Their approach requires that decisions regarding coverage for new technologies (and other limit-setting decisions) and their rationales must be publicly accessible. It also requires a “mechanism for challenge and dispute resolution re-
warding limit-setting decisions, including the opportunity for revising decisions in light of further evidence or arguments.\textsuperscript{51}

From a legal perspective the “accountability for reasonableness” factors map on to our understanding of basic requirements for procedural fairness; a duty to provide reasons and the ability to seek review in the general courts of decisions delegated from governments to administrative bodies.\textsuperscript{52} But presently we see a vast gap between the day-to-day reality of decision-making in Canada and the ideals expressed in the accountability for reasonableness framework. The principles that guide decision-making are not transparent or available for public scrutiny. Policy decisions are closed to public input. There are few formal avenues to appeal or review decisions apart from through the general courts. And the courts themselves have tended to be extremely deferential to governmental decision-making despite the failure of governments to demonstrate that fair processes have been followed. The overriding sentiment in Canadian health policy is that it is better to be secret rather than open. There is much that can be done to fix processes of decision-making within the present system without having to get involved in second-guessing decisions.

Presently, Charter review is not running amok over governmental decision-making within Medicare (although Chaoulli is a strong signal to governments that unless they run publicly-funded Medicare well the courts will not tolerate constraints on patients who wish to “go private”). But then, on the other hand, it is not doing much to improve the quality of decision-making either; it is not sending signals to decision-makers that they must be fair, open, and transparent. Rather the Supreme Court is signaling that it rarely wants to get involved in the allocation issue within social programs. Partly the problem is that a Charter challenge is simply not the best vehicle for mounting a challenge to governmental decisions limiting Medicare.

In terms of section 15, the problem in health care cases is that many patients could be classified as suffering from some form of pre-existing disability and in turn any denial could be readily portrayed as demeaning human dignity. Dignity is “inherently malleable. While it may be easier to determine when human dignity is demeaned, it will be more difficult to articulate why it is not.”\textsuperscript{53}

The courts have responded to the potential of this slippery slope in health care cases by responding formally to claims, and with Auton have largely closed the gate to claims that excluding certain disadvantaged groups from health insurance programs not intended to be universal is in and of itself discriminatory.

In terms of enhancing the process of decision-making, the most obvious Charter vehicle is section 7, which provides that “life, liberty and security of the person” is not to be deprived thereof except in accordance with the principles of fundamental justice, the latter offering a range of procedural protections that will vary with the degree of interest affected.\textsuperscript{54} The difficulty here is that in order to invoke the protections of section 7, a claimant must first overcome the threshold test of demonstrating that “life, liberty or security” is at issue. The courts have not historically read these provisions broadly and, prior to Chaoulli, had never found “life, liberty or security” of the person denied as a result of denial of publicly funded health care. The majority in Chaoulli were very clear to point out that the applicants were not seeking publicly-funded care but rather the right to purchase private insurance, and that given long waiting times in the public sector their rights to life and security, guaranteed by section 7, were violated by legislation that precluded this option. But this does raise the question of whether or not in a future case an applicant affected by long waiting times in the public sector may seek as a remedy, not the right to purchase private insurance, but the right to public funding for more timely treatment and/or fairer and more just processes for determining priorities on waiting lists.\textsuperscript{55}

**International experience, from the Netherlands, Oregon, New Zealand, and Israel, suggests that the project of determining an explicit list of principles to guide decision-making is often doomed to failure because of the intractability of ranking those principles.**

The courts have responded to the potential of this slippery slope in health care cases by responding formally to claims, and with Auton have largely closed the gate to claims that excluding certain disadvantaged groups from health insurance programs not intended to be universal is in and of itself discriminatory.

In terms of enhancing the process of decision-making, the most obvious Charter vehicle is section 7, which provides that “life, liberty and security of the person” is not to be deprived thereof except in accordance with the principles of fundamental justice, the latter offering a range of procedural protections that will vary with the degree of interest affected.\textsuperscript{54} The difficulty here is that in order to invoke the protections of section 7, a claimant must first overcome the threshold test of demonstrating that “life, liberty or security” is at issue. The courts have not historically read these provisions broadly and, prior to Chaoulli, had never found “life, liberty or security” of the person denied as a result of denial of publicly funded health care. The majority in Chaoulli were very clear to point out that the applicants were not seeking publicly-funded care but rather the right to purchase private insurance, and that given long waiting times in the public sector their rights to life and security, guaranteed by section 7, were violated by legislation that precluded this option. But this does raise the question of whether or not in a future case an applicant affected by long waiting times in the public sector may seek as a remedy, not the right to purchase private insurance, but the right to public funding for more timely treatment and/or fairer and more just processes for determining priorities on waiting lists.\textsuperscript{55}

**The Unfulfilled Potential of Administrative Law?**

There is no need, however, to invoke the constitution to achieve better decision-making in health care. There is potentially greater scope for law to require better decision-making using the far less glamorous field of administrative law. The threshold test for invoking procedural protections under administrative law is much lower than it is for a section 7 Charter challenge.\textsuperscript{56} On both procedural and substantive review, the courts are deferential to governmental decision-making in determining what is and is not publicly funded. Courts
demonstrate their deference by reviewing decisions on the standard of “patent unreASONableness.” This is the most deferential standard of review possible in administrative law, with the other possibilities being “reasonableness simpliciter” and “correctness” (the latter being the least deferential). Nevertheless, successful administrative law applications for judicial review of decisions not to fund medical treatments are rare.\(^{57}\)

Apart from substantive review of decision-making, there have also been few administrative law cases challenging the processes by which decisions are made.\(^{58}\) Partly this is because of a focus on Charter remedies – a desire to go for the big bang effect of overturning a governmental decision not to fund a treatment. Partly this is due to the deference given by the courts to polycentric decisions. And partly also this is because of the difficulty of latching on to decision-makers in the administrative law sense who may often be deeply embedded within Ministries of Health, or because the real decision-makers are often physicians, managing their own wait-lists, and have yet to be considered within the scope of judicial review. Notably in both Stein and Lalonde (see endnotes 57 and 58) there was an independent tribunal that made a specific decision and thus could be more readily reviewed in administrative law.

In this latter sense there are real barriers for bringing actions in judicial review of decisions determining the limits of publicly funded Medicare. The first recourse of patients seeking public funding for a treatment is to appeal to the Ministry. In that case, our investigations indicate that a Medical Director may often play a pivotal part in the decision of whether to deny or allow funding. Medical Directors are salaried physicians employed with the Ministry (and indeed every province employs several Medical Directors). A Freedom of Information and Protection of Privacy Act\(^ {59}\) application revealed that the Medical Directors from different provinces meet biannually through the auspices of the Inter-provincial Health Insurance Agreements Coordinating Committee.\(^ {60}\) The transcript of one of the meetings obtained suggests that provinces may pressure each other not to list new procedures and technologies because of the stress such action places on other provinces to fund treatments. Applications for access to more information about these meetings have been made to date, with no success on the grounds of the sensitive nature of such inter-provincial communications.\(^ {61}\)

Our research indicates that Medical Directors are important decision-makers, deeply embedded within the provincial ministries.\(^ {62}\) However, most citizens do not have any idea that these individuals are front-line decision-makers determining what services attract public funding and what do not. In this regard the most disconcerting aspect of the Cameron decision, which focused on funding of infertility treatments, was not the failure per se to publicly fund IVF treatment but rather the failure to demand fair processes for decision-making. The legislation in question provided for the establishment of an administrative tribunal; but none had ever been set-up, leaving critical and value-laden decisions about access in the hands of one medical professional within the Nova Scotia Ministry of Health.

The Canadian experience is in sharp contrast to that within the UK, (as more fully discussed by Chris Newdick in this volume) where there has been a shift from weak procedural scrutiny of decision-making to much more intense scrutiny.\(^ {63}\) Why has there been more of a shift towards tougher scrutiny in the UK as opposed to Canada? A partial answer is that administrative law has developed more robustly in the UK relative to Canada, where claimants are more likely to argue a Charter claim rather than an administrative law claim. But also I think the answer is that the UK has moved much more towards a culture of openness, transparency, and accountability in its health policy, in an effort to reinvigorate and improve the performance of the NHS. As a consequence, there are more obvious administrative decision-makers (tribunals at arms’ length from government) and more transparent decisions, making it in turn much easier for administrative law to be of relevance.

**Conclusion**

This paper has provided a relatively brief and high-level survey of the approach of Canadian courts to claims to access publicly-funded Medicare and/or to more timely treatment. Canadian courts have not, as is sometimes claimed, been proactive in expanding entitlements to publicly funded Medicare. In this regard courts seem to assume rationality in decision-making about what services are publicly funded and which are not. But there is mounting evidence that governmental decisions are guided by principles that no one objectively would consider meritorious (e.g. short-term political gains or long-held accommodations with the medical profession). Moreover, and more importantly, the processes of decision-making are neither fair nor transparent. One must then question the level of deference afforded by the courts to governmental decision-making and ask whether or not there are opportunities for courts to require that the principles upon which decisions are made are both transparent and fair.

Within the literature on prioritization, and in a number of jurisdictions, the emphasis has turned from attempting to articulate the principles that should determine what services are funded towards improving the fairness of the process of decision-making. In this re-

Colleen M. Flood
gard, claims in administrative law could potentially improve the transparency and fairness of decision-making. Notwithstanding, there have been very few administrative law challenges to the substance or the process of decision-making. Partly this may result from an undue focus on the Charter as a means of redress, but it is also due to the difficulty of identifying who makes decisions about the boundaries of public funding. Key decision-makers are buried deep within governmental bureaucracies and it is often unclear who is accountable for what. Canada has not embraced the kind of reforms that have been embraced in the UK and New Zealand which would clarify lines of accountability and improve the opportunities for administrative law to act as a check on decision-making.

In sharp relief to the court’s conservative approach to applications for better or timely access to publicly-funded Medicare is the recent Supreme Court of Canada’s decision in Chaoulli which upheld a right to buy private health insurance for “medically necessary” hospital and physician services. This decision overturned legislation in the province of Quebec prohibiting private health insurance. This case raises the prospect of future Charter challenges to provincial legislation designed to prevent the flourishing of a private sector. It seems that Canadian courts are destined to play a significant role in determining the future balance between public and private insurance in the Canadian health care system. Sadly, however, Canadian courts to date have staked out only a very limited role to help improve the internal fairness and operation of publicly-funded Medicare itself.

Acknowledgements

I would like to thank Sasha Kontic for her superb research assistance. I would also like to thank the Faculty of Law, University of Auckland for hosting me through my sabbatical year within which time this article was written. Thanks also to Sujit Choudhry, Lance Gable and Ted Hutchinson for their respective comments on the paper – all errors and omissions are mine.

References

14. Carolyn Tuohy, Mark Stabile and I am the principal investigator of a 3 year program of research exploring this issue. It’s a multi-disciplinary project with 10 investigators. To learn more about the project or read one of the 6 working papers please go to our website at <http://www.law.utoronto.ca/healthlaw/basket/> (last visited September 27, 2005).
18. Ibid. at para. 78. Justice Oliphant writes: “In my view, legislation that forces women to have to stand in line in an overburdened publicly-funded health care system and to have to wait for a therapeutic abortion, a procedure that provably must be performed in a timely manner, is a gross violation of the right of women to both liberty and security of the person as guaranteed by s. 7 of the Charter.”
19. This film depicts Quebec hospitals as both chaotic and squalid and the staff therein amenable to bribery. Such dramatization makes for a good story but certainly undermines trust in publicly-funded health care and fuels demands for privatization.


42. It is worthwhile to note that the _Auton_ and _Chaoulli_ decisions were heard at the same time although the _Auton_ decision was released much earlier than the _Chaoulli_ decision. It is also interesting, particularly given the importance of the facts to the eventual determination in each of these cases, that in both decisions the Supreme Court overturned the decisions of both the trial and appeal courts.

43. _Auton, supra_ note 39 at para. 37.

44. _Ibid._ at para. 38.

45. _Ibid._ at para. 43.


47. See the working papers at <http://www.law.utoronto.ca/health-law/basket/> (last visited September 27, 2005).


51. _Ibid._ at 57.


55. In _New Brunswick (Minister of Health and Community Services)_ v. _C._ (C.J.), (1999) 3 S.C.R. 46, a women was successful in obtaining public funding for legal aid to assist her in challenging a judicial order granting the minister custody of the applicant’s three children for an additional six months. See also S. H. Hartt and P. J. Monahan, “The Charter and Health Care: Guaranteeing Timely Access to Health Care for Canadians,” D. H. Howe Institute _Commentary_, No. 164 (May 2002) for an argument that in order to defend a section 7 challenge governments need to put in place wait time guarantees. Available at <http://www.charterhealth.ca/articles/edhowe_commentary.pdf> (last visited September 27, 2005).


57. The only successful judicial review claim before the courts with re-
guard to waiting times has been *Stein v. Quebec (Regie de l’Assurance-maladie)* (1999) QJ No. 2724 (S.C) (QL). In this case Mr. Stein waited months for surgery, even though his doctors warned his life was in danger if he was not operated on within four to eight weeks. He was successful before the Quebec Superior Court in overturning the Quebec health insurance board officials’ refusal to pay for his treatment in a New York hospital on the grounds that, given the facts of the case, the decision was patently unreasonable. The court was prepared to be very deferential to the Quebec authorities; however, even allowing for this very high standard of deference the court felt compelled to overturn the Board’s decision. Thus the courts will check the rationality of decisions about what is in and out of Medicare, but will generally not hold the government or other institutions to any higher standard.

58. The courts have also on at least one occasion sent a sharp message that constitutional values must be incorporated into decision-making. The Ontario Court of Appeal decision in 2001 of *Lalonde v. Commission de Restructuration des Services De Sante* (2001, 56 O.R. (3d) 505, involved a challenge to the Health Services Restructuring and Taskforce Commission’s decision to close the francophone Montfort Hospital. Both the Ontario Divisional Court and the Ontario Court of Appeal held that whilst s.15 of the Charter would provide no relief, applicants seeking to ensure the continuation of French language medical services to the francophone population in Ottawa-Carleton could succeed by relying on the unwritten constitutional principle of protection of minorities. The Ontario Court of Appeal found that the Commission, in failing to give serious weight and consideration to the importance of Montfort to the survival of the Franco-Ontario minority, had failed to exercise its public interest mandate as required by the fundamental principles of the Constitution. The court said at para. 180 that: “in determining the public interest, the Commission was required to have regard to the fundamental constitutional principle of respect for and protection of minorities...The Commission, however, viewed consideration of Montfort’s larger institutional role as beyond its mandate.”


60. Its predecessor was the Federal-Provincial/Territorial Coordinating Committee on Reciprocal Billing.

61. The Director of Access to Information at Health Canada has advised that although there are relevant documents, they will likely not be released to us as they fall within the exemption of information obtained “in confidence” from other governments, the disclosure of which could be “injurious to federal-provincial consultations” (*Access to Information Act*, R.S.C. 1985, c. A-1, s. 13(1)(c) and s. 14).


The amount allocated to publicly funded health care for 2005/06 in New Zealand, a small country of some four million people, is $NZ 9.68 billion, or 6.2% of GDP, an increase from the 5.7% of GDP in 2000/01. The Minister of Finance has recently signalled that spending in health and education has outpaced economic growth, and that the present rate of growth in health spending, which has grown at about 7% a year over the last decade, is unsustainable. Despite these big funding increases in recent years, the perception of New Zealanders is that the extra spending has made little difference, at least to hospital services and to people's ability to access treatment. In surveys, health emerges as a leading concern for New Zealanders. Their concern is apparently less about the quality of services, than about their ability to access treatment – whether they will be able to access timely health care when they or their family members need it.

Whatever proportion of GDP a country spends on health care, demand for health care outstrips supply. The gap is exacerbated by the increased effectiveness of medicine. People with chronic conditions are staying alive longer, being maintained on expensive treatments, in contrast to brief, acute episodes requiring terminal care in the past. New, better, and expensive medical technologies and pharmaceuticals continue to be developed and promoted, thereby creating increased expectations, in a more educated and consumerist population, of what health care can and should provide. There is more than ever a need to ration health care. But at the same time, people are less prepared to defer to decisions when they suspect access is being denied on financial grounds, and some are prepared to use the law to challenge them if necessary.

Joanna Manning, B.A., LL.B. (Hons.), M.Comp.L., is an Associate Professor at the Faculty of Law, The University of Auckland, where she teaches and has published widely on issues of medical law and ethics. She has degrees from the University of Auckland, and a Master of Comparative Law from the National Law Center, George Washington University, Washington DC. For 10 years she was the consumer representative on a national disciplinary tribunal with jurisdiction over physicians, and is a member of a national advisory committee to the New Zealand Minister of Health on matters of health care ethics. Ron Paterson, LL.B. (Hons.), BCL Oxon., has been New Zealand's Health and Disability Commissioner since March 2000. He has law degrees from Auckland and Oxford Universities, and was a visiting law professor at the Universities of Ottawa and British Columbia in the mid-1980s, before returning to The University of Auckland as senior lecturer in health care law (1986-99). Ron has worked in mental health, public health, and consumer protection policy roles at the Ministry of Health, including as Deputy Director-General, Safety & Regulation. He has taught and researched in the United States, as Fulbright Visiting Professor of Biomedical Ethics at Case Western Reserve University (1993) and Harkness Fellow in Health Care Policy at Georgetown University (1998-99).
New Zealand has adopted a number of specific innovations to improve access and quality for patients. The use of clinical guidelines to determine access to publicly funded health services, and the adoption of legally enforceable patients’ rights to safeguard quality, are of particular interest. We will describe the rising interest since the 1990s in priority setting in health care in New Zealand and assess the legal challenges to rationing decisions in recent years. In particular, we will focus on two significant challenges to rationing decisions, both relating to denial of dialysis treatment. We conclude that rationing is defensible under the New Zealand Bill of Rights Act, but if the result is indirect discrimination on the grounds, for example, of age or disability, governmental authorities must be prepared to demonstrate an objective, evidential basis for taking the prohibited ground into account to that extent.

Distinguishing Features of the NZ System

From a North American perspective, the New Zealand health system has two noteworthy features. First, the health-care system is one of universal coverage. In 2002 the Government funded 78% of total health expenditure from general taxation revenue. Secondary care is provided free of charge from publicly owned hospitals; primary care and prescription drugs are partially subsidized. Secondly, a public “accident compensation” insurance scheme covers the cost of rehabilitative medical care and earnings-related compensation for victims of “treatment injury” or work, traffic, and other unintentional injuries. Moderate lump-sum compensation is also available, and tort claims for damages for negligence are barred if treatment injury insurance cover is available. The net result is that health professionals are the beneficiaries of a system that looks to the state to compensate victims of medical negligence. Doctors pay modest professional indemnity levies, and are not required to contribute to the cost of state funding for treatment injuries covered by the accident compensation scheme. The universal coverage system promotes equity of access, and in theory the bar on medical negligence claims removes a major barrier to quality improvement.

Legislative Framework for Rationing

The New Zealand legislature has been remarkably open in recognizing the need to ration access to publicly funded health care – even though politicians and health officials prefer to speak of “prioritizing” access. Resource constraints are expressly recognized in the funding framework, the New Zealand Public Health and Disability Act 2000: the system’s objectives – a population health focus, equity, community participation, and access to appropriate, effective, and timely services – are to be pursued “to the extent that they are reasonably achievable within the funding provided.” Echoes of this approach are found in the Injury Prevention, Rehabilitation, and Compensation Act 2001, which includes “failure to provide treatment, or to provide treatment in a timely manner” within the definition of treatment injury covered by the accident compensation scheme – but then excludes “personal injury that is solely attributable to a resource allocation decision.”

And even in the Code of Patients’ Rights, the various rights (including Right 4, to services of an appropriate standard) are qualified by a defense available to providers who show they took “reasonable actions in the circumstances to give effect to the rights,” with “the circumstances” defined to include “the provider’s resource constraints.”

The explicit authorization of rationing is circumscribed by anti-discrimination legislation: the Human Rights Act 1993 (NZ). Section 44(1)(b) makes it unlawful for a provider of services to the public “to treat any...person less favourably in connection with the provision of those...services than would otherwise be the case, by reason of any of the prohibited grounds of discrimination.” The prohibited grounds include “age” and “disability.”

The negative right (not to be denied health services by reason of unlawful discrimination) is not matched by an affirmative right to access health care. There are only two examples of a “duty to treat” in a health statute: § 66 of the Mental Health (Compulsory Assessment and Treatment) Act 1992 (NZ) and § 51 of the Intellectual Disability (Compulsory Care and Rehabilitation) Act 2003 (NZ) state that every patient/care recipient subject to the compulsory regime is “entitled to medical treatment and other health care appropriate to his or her condition.” It is obviously incumbent on the state, having deprived an individual of liberty for the purpose of compulsory treatment or care, to ensure that appropriate treatment and care is provided.

Nor does the Code of Patients’ Rights contain a right to access services. The statutory parameters for the Code confine it to quality of service issues. However, the line between quality and quantity of service may not be as clear as has generally been assumed, and parts of the Code (e.g., Right 4(3), “to have services provided in a manner consistent with” the consumer’s needs, and Right 7(8), “to express a preference as to who will provide services and to have that preference met, where practicable”) come close to giving a right of access.

The New Zealand Bill of Rights Act 1990 contains no affirmation of the right to health care, in contrast to documents such as the Constitution of South Africa (§27(1)). New Zealand is a signatory to international instruments such as the International Convention on
Economic, Social and Cultural Rights; however, the rights affirmed (such as the right to health, in article 12) have not been directly incorporated into domestic law – although the courts have hinted that they may be mandatory relevant considerations for decision-mak-

ers exercising a public function (e.g., allocating public funding), thereby opening up the possibility of judicial review absent such consideration.14 The Human Rights Amendment Act 2001 (NZ) gave the Human Rights Commission responsibility for the development of a national plan of action for the promotion and protection of human rights in New Zealand, and its inaugural plan sets as a priority for action to “initiate a structured public discussion about a human right to health.” The tentativeness of this approach by the Commission underlines New Zealand’s wariness in relation to legislated rights to access to health care.

A Summary of Priority Setting in New Zealand

For over a decade there has been a rise in interest in New Zealand, as elsewhere, in developing and using explicit criteria for rationing health care. Indeed, New Zealand has been considered by some international experts to be among the “leading bunch” in the move to explicit rationing:15 Historically, New Zealand’s experience matched that elsewhere in the developed world. Rationing of health services was implicit.16 Clinicians made decisions about resource utilization at the individual patient level in an ad hoc manner on the basis of clinical judgment. The fact that decisions were driven by resource constraints was seldom acknowledged, and rarely suspected by patients.

This process of “muddling through and evading responsibility”17 became unsustainable in the early 1990s in New Zealand, when an incoming conservative government moved to reform the public health system to introduce a form of regulated competition in line with then current international trends.18 The main planks of the reforms – the purchaser-provider split and re-structuring of what were seen as inefficient and financially voracious public hospitals into companies with a profit-making imperative19 – aimed to introduce more competition into the funding and provision of health care, thereby improving efficiency.20 The fact of finite resources was explicitly acknowledged in the purpose of the statute implementing the reforms. The social purpose of “secur[ing] the best health, the best care or support for those in need of those services, and the greatest independence for people with disabilities” was to be pursued only so far as “reasonably achievable within the amount of funding provided.”21 There was no guarantee of access to services, the purpose of public funding and provision instead being limited to “facilitat[ing] access to personal health services and to disability services.”22

An important plank of the reforms, designed to alleviate public anxiety, was to define a list of “core health services,” to which there would be a level of guaranteed access within the constraints of funding. Funders would be required to purchase these core services as a matter of priority, and they would be available without charge or at affordable cost in the public sector within a reasonable time. Responsibility for defining the core was delegated to a National Advisory Committee on Core Health and Disability Support Services set up in 1992.23 A key task of the committee was to advance public debate and understanding of limited health-care resources and the need to make choices. Thus, the fact of scarcity, the need for prioritizing services, and the identity of those responsible for rationing decisions now became publicly visible. The public was to be involved in making decisions about which public health services should receive public funding, or at least be aware that such decisions were being made.

Within two years, the Committee had rejected as neither fair nor workable a simple “Oregon”-type list of included/excluded services to which New Zealanders should have free access. Instead, it shifted its emphasis to developing mechanisms for prioritizing patients within services rather than methods for prioritizing between services. It embarked on the systematic evaluation of services where there was high use, high cost, or less about the quality of services, than about their ability to access treatment.

In surveys, health emerges as a leading concern for New Zealanders. Their concern is apparently less about the quality of services, than about their ability to access treatment.
REFERENCES, which included a mixture of expert professional and community representatives. From 1996, the Committee persuaded the government to fund an extended guidelines program, accelerating the process of guidelines development and implementation. Guidelines for access to dialysis treatment for end-stage renal failure were at the heart of the South Auckland Health case, a highly controversial and successfully challenged decision not to offer treatment to an individual patient, and revised access guidelines were the subject of judicial scrutiny in the Shortland case three years later. In 1999 a new center-left government, uncomfortable with a health system still based on a model introduced to implement competitive contracting for health services, embarked on far-reaching restructuring of the sector.

The Committee also advocated replacing waiting lists for non-urgent, routine procedures with booking systems designed to provide greater equity of access to services. Patients are no longer allocated to public hospital waiting lists; instead, they are booked for a first specialist appointment for assessment within six months, and then either given a “booked” date for their operations within six months or, if they do not meet the agreed criteria, are referred back to their general practitioner to be kept under “active review” (i.e., for re-referral if their condition deteriorates or if hospital capacity for so-called “elective” procedures increases). Key elements of the booking systems are nationally consistent referral guidelines developed by general practitioners and hospital physicians (Access Criteria for First Assessment or ACA), and agreed priority criteria for access to elective surgical procedures (Clinical Priority Assessment Criteria or CPAC). An identified advantage of the booking system is that it enables the gap between clinical need and ability to benefit and the levels of available funding to be identified. Publicly available information on levels of unmet need provides more opportunities for advocates and the public to compare access in different regions and make the case for more funding or improved ways of delivering care.

Access to publicly funded pharmaceuticals has been regulated via a pharmaceutical management agency, Pharmac, established to “secure for eligible people in need of pharmaceuticals, the best health outcomes that are reasonably achievable from pharmaceutical treatment and from within the amount of funding provided.” Pharmac manages the Pharmaceutical Schedule, which is a list, updated monthly, of subsidized prescription drugs and related products. It makes the final decision on subsidy levels and prescribing guidelines and conditions, by balancing evidence of effectiveness with cost within a fixed pharmaceutical budget. In making prioritization decisions it receives independent expert medical input from a Pharmacology and Therapeutics Advisory Committee. It has developed considerable expertise in using explicit processes, including cost-utility analysis, to inform funding decisions. Compared to many overseas countries, which have experienced significant growth in pharmaceutical expenditure, Pharmac has been highly successful in containing pharmaceutical costs while increasing access to medicines.

Publicly funded health services have, since 1993, been purchased by regional purchasers. In 1998 the four regional purchasing agents were amalgamated into a single funder, the Health Funding Authority (HFA). This provided the opportunity for renewed effort to improve priority-setting processes in purchasing services. After some public consultation, the HFA did further work developing an explicit process for between-service priority-setting, based on explicit values combined with an analytical methodology to assess effectiveness and cost. The HFA used the principles developed by the Core Services Committee in 1993, which led to the inclusion of a fifth principle: the need to improve Maori health. During 1999/2000 the HFA applied the process to determine the relative priority of services in allocating new money for the following year, which accounted for approximately 1% of the health budget. In 1999 a new center-left government, uncomfortable with a health system still based on a model introduced to implement competitive contracting for health services, embarked on far-reaching restructuring of the sector. The single funding authority was abolished. The centerpiece of these reforms were twenty-one District Health Boards (DHBs) responsible for funding from allocated government funds almost all publicly funded health and disability support services in their regions.

DHBs fund and, through their hospitals, provide much of secondary care, as well as being responsible for funding primary care services through contracts with community and primary care providers. One of the purposes of the reforms was to provide for greater public input into decision-making by providing for a majority of members of Boards of DHBs to be elected. Purchasing decisions are now made locally within national policy frameworks rather than nationally. DHBs are required to consult with their local populations, to sur-
vey and assess their health needs to inform their spending priorities, and to develop their own priority-setting processes. The move to twenty-one separate DHBs raises the specter of local variations in access to services and new health interventions. The fragmentation of expertise and the duplication of effort across twenty-one DHBs in developing prioritization processes appear to have left some, especially smaller DHBs, with a lack of capacity for carrying out the task. As a result, “explicit priority-setting is increasingly being seen as a complex task that is beyond their capacity”—something to be undertaken at a regional or national level. The evidence from the first two years of the new structure suggests that, although required to carry out health needs assessments of their local populations and determine priorities based on them, the DHBs’ ability to allocate resources based on local preferences has been extremely limited owing to a tight financial environment and predetermined health priorities in national strategies, which DHBs’ plans are required to reflect. Almost all DHBs saw prioritization as applying to new funding, with 99% of their budgets determined by historical patterns of resource allocation. A period of hiatus seems to have occurred during which momentum in developing prioritization processes has slowed. Co-operation and co-ordination between DHBs in priority-setting may be part of the solution. There are signs of this developing, for example in moves to develop a collaborative process between DHBs for making decisions about adopting new health interventions.

Until implementation of the 2000 reforms, the single funder, the HFA, was the principal decision-maker relating to funding new health interventions. Its prioritization process included new interventions. On occasion the HFA and later the Ministry of Health, which initially assumed the function under the new structure, were obliged to fund new technologies because of political decisions made by government. With DHBs now playing the critical role in most decisions about funding of new interventions, there is concern about the absence of a systematic process within and between DHBs for assessing new interventions.

In recent work the National Health Committee has provided a fascinating snapshot of current decision-making processes within DHBs about the adoption of new interventions. Clinicians are currently the key players in decisions within hospitals. Decisions tend to be made within a clinical specialty, and as a result are usually based on historical patterns of service, rather than being seen or treated as related to wider prioritization processes. As a result there is a potential for decisions within hospitals to be inconsistent with DHB priorities and to increase health inequalities. A wide variety of formal and informal factors influence decisions. Some of the formal processes set up in hospitals to improve consistency and quality in decision-making are not particularly robust or are easy to avoid by clinicians. The “domino effect” was noted— if one DHB decides to fund a new clinical intervention, others feel they should follow suit. It is also inefficient for each DHB, especially smaller ones, to duplicate an extensive evaluation of the same new interventions. Lack of consistency between DHBs is a constant theme. There is little use of health technology assessment information in formal decision-making processes, although most DHBs require some, albeit limited, evidence of effectiveness. The Committee concluded that the problems are wider than just the lack of an adequate evidence base for decision-making. There is a general lack of capacity and capability within DHBs to undertake assessment of evidence. The Committee rejected the establishment of a centralized decision-making institution to make all or most decisions about the introduction of new technologies and to issue directives to DHBs as contrary to the DHB model, but it identified the need for better collaborative processes between DHBs, particularly for high-cost interventions or complex decisions. It endorsed the former HFAs principles-based approach. To improve decisions about new health interventions, it has recommended to the Minister of Health that priority should be given to developing robust decision-making processes and encouraging and monitoring their adoption throughout the sector. The second area of priority is improving the capacity and capability to access and assess evidence and information about new interventions. Although it proposed some possible solutions, the Committee considered that DHBs are in the best position to consider these options and decide collaboratively the best solution to support their requirements for evidence and information.

Commentators observe that “{e}fforts to prioritize funding explicitly have led to positive changes. There is greater openness about finite resources and increasing awareness of the costs of different healthcare practices. Hospital efficiency has improved, as shown by reductions in length of hospital stay and increased throughput. Booking systems and referral guidelines are improving equity of access to elective hospital services.” But the new arrangements have not been without their critics. Some consumer advocates and clinicians argue that the improvements are illusory, and that patients are being denied potentially beneficial treatments on the basis of thresholds reflecting financial capacity rather than clinical need.
Legal Challenges

There have been relatively few legal challenges to decisions about the public funding and provision of health and disability services. Most of those challenges have involved judicial review of funding decisions, sometimes referred to as “meso” allocation of health resources. Examples include a number of challenges to funding decisions related to pharmaceuticals, provincial hospitals, rest homes, and patient subsidies for general practitioner services.

Legal challenges to rationing of health care at the individual patient level, sometimes referred to as “micro” allocation of health resources, have been even rarer. Two such cases, both involving the withholding of dialysis from patients with end-stage renal failure, galvanized public concern about perceived dwindling access to publicly funded health care. Both occurred during the 1990s, and were attended by intense public and media interest. The first, the South Auckland Health case, which did not proceed to the courts, highlights the potential for the Human Rights Act to be invoked by patients who allege that the failure to provide a treatment is discriminatory. The second, Shortland v. Northland Health Ltd, which resulted in two High Court rulings and a Court of Appeal decision, illustrates the difficulties in seeking judicial review of clinical decision-making.

In 1995, seventy-six-year-old James McKeown was initially denied dialysis treatment for his end-stage renal failure, by renal physicians at Middlemore Hospital (part of South Auckland Health). His family laid a complaint of age discrimination with the Human Rights Commission alleging that, in applying a guideline that “in usual circumstances, persons over 75 years are not likely to be accepted onto a…dialysis programme,” South Auckland Health had breached the Human Rights Act 1993.

The ultimate reversal of the decision to withhold dialysis may have been partly attributable to public pressure, fueled by a media campaign and complicated by complaints that the Minister of Health had breached the patient’s privacy in revealing confidential details of his co-morbidities on national television. However, the decisive factor appears to have been the allegation of age discrimination. South Auckland Health – understandably wary of the opprobrium attaching to a finding of unlawful discrimination – ordered its physicians to reassess the seventy-six-year-old patient ignoring his age, and dialysis treatment was promptly provided. Mr. McKeown enjoyed another eighteen months’ life.

In the second case, the hospital did not back down, and the case ended up in court. It concerned Rau Williams, a sixty-three-year-old Maori man with diabetes who was in end-stage renal failure. In June 1997 he was admitted to Northland Health’s Whangarei Hospital and placed on interim dialysis to enable his suitability for acceptance on to the hospital’s renal replacement program to be assessed – specifically his suitability for long-term, home-based peritoneal dialysis (CAPD).

An interdisciplinary team considered his suitability clinically and socially over a ten-week period, during which time five or six meetings were held, attended by family members. Northland Health then advised the family of its decision to discontinue dialysis. The decision was made in part by applying an evidence-based access guideline for entry to the program developed by the regional health authority responsible for funding the treatment. The guideline allocated the resource on the basis of ability to benefit from treatment, with the proviso that in the usual case a prognosis of two years’ life or “benefit” was required for automatic admission to the program. Williams’ prognosis was that continued dialysis would have extended his life by about one year. Of those patients not expected to derive the two-year benefit, some also fell into Group A of the guideline, which listed factors that in isolation were likely to result in a decision of unsuitability, and for whom admission on to the program was then exceptional.

The relevant section of the guideline read:

**Group A**

Factors which in isolation are likely to determine that an individual is not suitable for treatment of End Stage Renal Failure

- (CNS / Mental Function)
  - Dementia (moderate to severe), very low IQ, a disabling psychiatric disorder which is unlikely to respond to further therapy, previous major stroke with persisting severe functional disability.
  - Basis: There must be the ability to co-operate with active therapy.

Williams had moderate dementia, a complication of the diabetes that had caused his renal failure and need for dialysis. Attempts to teach him to perform CAPD during the assessment period were unsuccessful. The treating renal physician concluded that he would be incapable of living independently and of performing

---

Even if the hospital’s budget was not capped, it is hard to imagine that admitting an additional patient would not impact on the regional or national dialysis budget.
any form of home dialysis, though family members indicated their willingness to supervise. He was considered unable to “co-operate with active therapy” by reason of his dementia, in terms of Group A of the guidelines, and hence considered unsuitable for entry on to the program. As a result of the family’s protests, opinions were obtained from five renal physicians from major centers around the country, all of whom concurred that discontinuing dialysis in the circumstances was appropriate and consistent with national practice.

The family challenged the withdrawal of dialysis on judicial review, seeking an interim order requiring Northland Health to continue or resume dialysis until a full judicial review proceeding could be heard. The patient’s evidence was that he did not wish to die; he enjoyed some quality of life on dialysis, including pleasure from seeing his family. The same day that the second High Court decision was announced, an urgent appeal was brought, heard, and determined by the Court of Appeal. The challenges were to no avail, and Rau Williams died of kidney failure early on the morning after the Court of Appeal’s decision.

On the first application to the High Court it was argued that the Health and Disability Services Act 1993 (NZ) imposed a duty on the doctors. Justice Salmon did not accept that the Act imposed a statutory duty to treat on Northland Health: “the general obligations ... set out in the Act are subject to clinical judgment” and are qualified by reference to what is “reasonably achievable within the amount of funding provided.” Justice Salmon noted that “[i]t is totally inappropriate for the Court to attempt to direct a doctor as to what treatment should be given to a patient.” The reluctance to impose a duty to treat, in the absence of clear statutory language, reflects the traditional view that clinical decision-making is the domain of clinicians, not courts. The principle reflects two concerns: first, a recognition that clinical judgment is beyond the expertise of courts; and secondly, that a court should not make orders with consequences for the utilization of scarce resources since it lacks knowledge of the competing claims to those resources.

Justice Salmon concluded by observing that it was unclear whether the renal physician’s decision to cease dialysis was made on the basis of a clinical judgment as to Mr. Williams’ best interests, on an assessment of where scarce resources should best be used, or on a simple application of the guidelines. He added that “to allow the Guidelines to dictate the result would be a reviewable error of law,” whereas a conclusion guided by the policy would be lawful. The hospital needed to clarify the reasons.

The applicant brought a second application two weeks later, raising various fresh grounds. Somewhat predictably, one alleged illegal fettering of discretion by slavish application of the guidelines without consideration of the individual circumstances of the patient. But this ground was unsustainable in the face of further evidence from the treating physician that the guidelines were used as “exactly that: a guide; a framework within which I had to make a clinical judgment.” The Court was satisfied that the guidelines were treated properly as a guide and did not dictate the final decision. Accepting that there was a resources element in the decision-making process when applying the guidelines, Justice Salmon considered that the decision was nevertheless a clinical decision, even though one of the factors taken into account is the availability of resources. “The resources issue, if it played any part in the decision, was a minor component.”

If this statement stretches credulity, it pales in comparison with the Court of Appeal’s later denial of any resource element in the decision (discussed further below). Justice Salmon was in any event prepared to accept the lawfulness of the decision to the extent that it was made pursuant to guidelines premised on an attempt to allocate a scarce resource fairly.

The value of the guidelines...is that they assist doctors to allocate scarce resources in a principled fashion...[A] properly made clinical decision which includes as an element of that decision considerations of resource allocation constitutes lawful grounds to decline treatment....Any other conclusion, apart from the difficulties that it would create for the medical profession, would create unacceptable inequities in relation to other patients competing for those resources.

This is a welcome statement. The High Court took a causation route around the “right to life” protection in New Zealand’s Bill of Rights Act 1990: “Mr Williams was ‘irretrievably doomed.’...[T]he decision, made as it was in good faith and in accordance with the doctor’s best clinical judgment, was not a decision to deprive of life, but rather one to let life take its natural course.”

In the Court of Appeal, the applicant’s first argument was that Northland Health’s refusal to provide dialysis amounted to breach without lawful excuse of its duty in criminal law to provide the necessaries of life to Rau Williams. It was without lawful excuse, because it had not complied with the requirements of “good medical practice” as defined by the High Court in Auckland Area Health Board v. Attorney-General. That was a case about the legality of withdrawing artificial ventilation from a patient with extreme Guillain Barré syndrome, where the family agreed that the proposed course of action was in the patient’s best interests. Justice Thomas had held that the life-support system could
be withdrawn without medical and nursing staff attracting criminal responsibility for the ensuing death of the patient if the discontinuance was in accordance with "good medical practice," defined as comprising (1) a decision in good faith that withdrawal of the life-support system was in the best interests of the patient; (2) conformity with prevailing medical standards and with practices, procedures, and traditions commanding general approval within the medical profession; (3) consultation with appropriate medical specialists and the medical profession’s recognised ethical body; and (4) the fully informed consent of the family. In *Shortland*, the third and fourth criteria had not been complied with. Secondly, it was argued again that Northland Health’s decision was in breach of Williams’ right to life.

As for the duty to provide the necessaries of life, the Court held that the clinical circumstances provided a “lawful excuse” for omitting to provide dialysis. The Court referred to the guidelines, recording that, because demand for renal dialysis had steadily outstripped the medical resources available, the guidelines were developed to assist providers to ensure comparable services across the region and then missed opportunity to educate a public (whose attention was focused on a particular case) about the inescapable fact of finite resources and the consequential need to ration, and to provide judicial support for an explicit and principled means of doing so.

The Court considered the four *Auckland Area Health Board* criteria. The first, a *bona fide* decision that withdrawal of life-supporting treatment is in the patient’s best interests, was the subject of a “proper” concession, for “there [could] be no doubt that the decisions in issue were made in good faith in the belief that they were in the best interests of Mr. Williams.” The matter was far from inarguable. Respect for autonomy requires that a competent patient is usually considered the best judge of his own best interests. New Zealand has a Code of Rights, then in force. It received no mention in the judgment. Consistently with common law principles, the Code provides for a rebuttable presumption of competence. There was no consideration of the fundamental issue of Mr. Williams’ competence. Although he was probably incompetent as he neared death at the time of the Court’s decision, he may well have been competent some months earlier when the decision to cease dialysis was made. Then and later he had indicated that the quality of life achieved through treatment was sufficient for him to want to go on living. If competent, the denial of treatment was a decision imposed on him by doctors, sanctioned by the Court, that his desire to continue to live was misguided. It was better for him to die now of kidney failure rather than live longer on dialysis – a striking example of medical paternalism approved by the Court, with fatal consequences.

To be fair, the Court must be taken to have assumed Williams was incompetent to decide, for why else the concentration on his best interests without reference to his wishes? If so, and the proposed course of treatment was palliative care not involving dialysis, Rights 7(3) and (4) of the Code of Rights were relevant. Neither was considered. Essentially a codification of common law principles, Right 7(3) provides that a patient with diminished competence retains the right to make informed choices and give informed consent to the extent appropriate to his or her level of competence. And Right 7(4) provides for treatment to be able to be provided to an incompetent adult if in his best interests and if various conditions are met. One such condition is that the proposed treatment is not inconsistent with the patient’s known wishes. If not a resource-based decision, both rights were breached, as Williams had indicated his wish to be treated. Arguably, he retained sufficient competence to give that indication and have it respected or, if not, this was a decision taken for him to provide one course of treatment (palliative care) rather than another (dialysis), in the face of his recently expressed preference for the latter. Absent any resource constraint, there was no reason not to accede to his preference. His inability to self-administer home-based CAPD could have been overcome by providing a dedi-
cated nurse to administer it, or by providing him with haemodialysis treatment in hospital. And, in requiring only a good-faith decision believed to be in the patient’s best interests, the Court accepted a very low standard of review, especially in a case where a life was at stake.90 What of an irresponsible or negligent decision? Even in Bland, the House of Lords measured decisions to withdraw life support against the Bolam test,91 and the law had moved somewhat since.92 It is now accepted that once its jurisdiction is invoked, the court is obliged to determine for itself the best interests of incompetent patients, and that a broad assessment of medical, emotional, social, and all other welfare issues is required, because “best interests” is wider than medical considerations.93

The Court dispensed with the need for satisfaction of the third and fourth Auckland Area Health Board criteria – ethics committee consultation and approval, and the fully informed consent of the family. While formal consultation with an ethics body might have been necessary on the facts in Auckland Area Health Board, it was not in this case, “which did not raise significant ethical issues.” “The issues arising were essentially ones of clinical judgment, not ethics.”94 Whatever the nature of the decision, resource-driven or best interests, this claim is unconvincing. Even if a best-interests case, in determining Williams’ best interests his physician had to balance length of life against quality of life if treatment were given, compared to that achieved by palliative care. That involves highly subjective assessments of what is an acceptable quality of life. And if, as was denied, this occurred in the context of insufficient resources, requiring comparison with what could be achieved in other patients, it is trite to observe that the process is not a purely technical process, but is highly contested with a heavily moral component.95 The Court did suggest that ethical consideration was built into development of the guidelines, so further consideration was not required. This is a fair point, but only really a satisfactory answer if this was not simply a best-interests case and the resource dimension was acknowledged.

The Court also held that consent from the family was not required for the provision of treatment to, or the withdrawal/withholding of treatment from, an incompetent patient; only reasonable consultation and taking account of views.96 Otherwise the family would be given the power to require the treatment to be given or continued, a proposition “appropriate in the context of proposed withdrawal of a life-support system,” but one the law could not countenance...in the different situation of a decision to put a patient on long-term dialysis.97 The Court never spelt out the reason for the difference between the two cases. One distinction was that in Auckland Area Health Board the doctors with family agreement were seeking the Court’s sanction to withdraw life support, while Mr. Williams and his family were insisting on continued provision of treatment. It may be that the Court realized the potential difficulties in permitting families to dictate the provision of expensive, life-prolonging therapies. But the Court denied that any key difference existed. If there is no meaningful difference between the two cases, there is no justification for applying different principles.

Finally, the Court rejected the argument that the withholding of dialysis constituted a breach of Mr. Williams’ right not to be deprived of life. Taking what has been referred to as a definitional approach to fundamental rights,98 the Court considered that Northland Health’s actions of refusing to provide dialysis treatment would not “deprive” Mr. Williams of his life in terms of § 8.99 The reasoning was as follows. The careful process of assessment on the basis of the guidelines, which included reasonable consultation of the patient and family, amounted to a “lawful excuse” in terms of the criminal provision imposing the duty to provide the necessities of life. “Equally...it could not be said that [Northland Health’s] actions...would deprive” Mr. Williams of his life100 – something of a logical non sequitur. It may have made for more satisfying reasoning to construe § 8 as a negative right, not requiring the taking of positive action to protect life, at least in this case of a principled allocation decision, or to concentrate the real debate around the qualification to the right101 or the balancing considerations in the “justified limitations” provision.102 After all, it was accepted that if Mr. Williams was treated, his life would be extended.

So, what to make of the Shortland decision? Traditionally, courts have been extremely reluctant to review allocation decisions, as being not justiciable or beyond their competence,103 despite urgings to scrutinize the process of decision-making more closely to ensure procedural fairness and assist in addressing what has been called a “legitimacy gap.”104 In Shortland, clinicians had...
abandoned implicit rationing and attempted to specify explicitly the criteria for deciding whom to offer treatment. Certainly, there was more reason than usual for deference to support their efforts. Ironically, the Court’s reasoning “converted” a decision on resource allocation into a decision about clinical judgment, thus obscuring the process of rationing behind a “veil” of clinical judgement, just as if it was a case of implicit rationing. The absence of an internal appeals process from a decision applying the guideline to an individual patient – the third of Daniels and Sabin’s conditions of “accountability for reasonableness” – meant that it fell to the Court to satisfy this requirement via the “enforcement condition” (judicial review). The Court’s extremely limited scrutiny of the guidelines and the non-treatment decision, and the failure even to acknowledge the true rationing basis of both, prevented explicit consideration of the extent to which Daniels and Sabin’s publicity and reasonableness conditions were met. Arguably, both were satisfied (subject to the discrimination concern we discuss next). It does seem, however, that the process of developing the guidelines and the careful assessment in applying them were the reason the Court supported the decision. The development and dissemination of the guidelines, and the transparency of the assessment process, which included a reasonable opportunity for consultation by the family, were indicative of procedural fairness. The guidelines accorded with renal physicians’ practice throughout New Zealand and there was an overwhelming consensus of clinical opinion that Williams should not be offered long-term dialysis, so no hint of Wednesbury unreasonableness arose.

**Discrimination and the Human Rights Act 1993**

Regrettably, because the argument was never advanced, the Court in *Shortland* did not consider the potentially powerful argument that the refusal to offer dialysis to Rau Williams amounted to unlawful disability discrimination. New Zealand has two statutes protecting the right to be free from discrimination: the New Zealand Bill of Rights Act 1990 and the Human Rights Act 1993 (NZ). The former contains a set of rights, primarily civil and political, which regulate and limit the powers of executive government and public actors. One of the rights, § 19 of the Bill of Rights Act, is the right to be “free from discrimination on the grounds in the Human Rights Act 1993 (NZ).” The Human Rights Act 1993, despite its name, is in large part an anti-discrimination statute. Written primarily with the private sector in mind, it was intended to protect citizens from discrimination at the hands of fellow citizens, though it applies also to government acting as an ordinary person, such as landlord or employer. Whereas the Bill of Rights Act applies across the whole range of governmental activity, the Human Rights Act applies in certain limited areas, such as to the supply of goods and services, employment, accommodation, and partnerships. In relation to the supply of goods, facilities, or services to the public or any section of it, the Act states that it is unlawful for any person (a) to refuse or fail on demand to provide any other person with those goods, facilities, or services; or (b) to treat any other person less favorably in connection with the provision of those goods, facilities, or services than would otherwise be the case “by reason of any of the prohibited grounds of discrimination.” The prohibited grounds are: sex; marital status; religious and ethical belief; color; race; ethnic or national origins; disability; age; political opinion; employment status; family status; and sexual orientation. “Disability” is defined to include “intellectual or psychological impairment.”

Since the reason given for declining to provide dialysis was that Williams suffered from moderate dementia, as a result of which he was unable to self-administer CAPD, there was at least a *prima facie* case of disability discrimination to answer. And, as the decision was made by applying the guidelines, the ultimate issue was whether the guidelines themselves were discriminatory. Certainly, on its face, Group A of the guidelines was very broad. The same rationale that underpinned the “mental function” aspect of the guideline applied to Williams could be applied also to withhold treatment from a blind or an intellectually disabled patient, albeit with a similar prognosis. Such a result would surely be thought unacceptable. It may be that the refusal to provide dialysis to Williams by reason of his disability would have been considered justified by a special exception in relation to disability in the Act, pursuant to which it is a defense to show that an individual’s disability requires services to be provided “in a special manner” and that the provider “cannot reasonably be required to provide them in that special manner.” The issue would have been the lengths to which a health services provider should reasonably be expected to go to accommodate the needs of an individual patient with a disability to enable him to access a life-prolonging treatment.

Although it may have been key in *Shortland*, the disability exception is likely to have limited application in other guidelines embodying priority criteria. It is relevant only to the disability ground, and in any event not all health services need to be provided in a special manner to patients with disabilities. What of patients with co-morbidities, as a result of which their prognosis is poorer, such that they would be unable to demonstrate the necessary “benefit” to qualify for treatment under a
Disproportionate numbers of Maori patients New Zealanders is poorer than non-Maori, non-Pacific New Zealanders is the major reason for the disparity in life expectancy between Maori and non-Maori.119 Chronic disease is the major reason for the disparity in life expectancy between Maori and non-Maori, non-Pacific New Zealanders. Diabetes makes a sizeable contribution.120 Disproportionate numbers of Maori patients

The New Zealand legislature has been extremely reticent in legislating guarantees of access to health services will require dialysis. Yet Maori patients will disproportionately fail to qualify for it because of their relatively poorer health status. It is also recognized that clinicians frequently take into account social factors when forced to decide priority between patients. Some, such as ability to work, care of dependents, living independently, appear to invoke the prohibited grounds of discrimination on family or employment status. These factors were assigned a 10% weighting in priority criteria developed by the National Health Committee, for example.

The allocation mechanism in Shortland was the ability to benefit for those in need of treatment – a maximization or efficiency measure. The public can legitimately expect the most efficient use of resources paid for out of the public purse. But allocation based on cost-effectiveness is acknowledged to have potentially inequitable and socially unacceptable consequences. The potential exists for an unfair distribution of treatment of groups of patients defined by characteristics such as age, socio-economic status, race or ethnicity, or levels of disability. Although it is their lesser ability to benefit, rather than the prohibited characteristic, that is the reason for the decision to withhold treatment, conduct not intended to be directly discriminatory on prohibited grounds but having such an effect amounts to unlawful indirect discrimination under New Zealand’s Human Rights Act.

Taking into account factors such as age and comorbidities (disability) to inform ability to benefit to decide priorities between patients (whether in clinical assessments or in priority criteria) is prima facie discriminatory. The Human Rights Act does permit defences of “genuine justification” and “good reason” to be made out. But before the legality of rationing practices such as that in Shortland was tested, Parliament amended the Human Rights Act 1993. As a result, the issue now falls to be determined, at least in so far as publicly funded health care is concerned, under the New Zealand Bill of Rights Act 1990, rather than the Human Rights Act.

The purpose of the Amendment was to ensure that the Bill of Rights Act, rather than the Human Rights Act, should state the general anti-discrimination principle to which government in its governmental function, as well as anyone acting pursuant to a statutory or public function, should adhere. Thus, any potentially discriminatory actions are to be assessed against the Bill of Rights Act standard, which is more flexible. Section 19 of the Bill of Rights Act refers to “discrimination” rather than “different treatment,” as in the Human Rights Act. “Discrimination” was thought to have greater potential for definition to accommodate situations where different treatment is justified for good reason. But, most importantly, the Bill of Rights Act 1990 contains § 5, so that there can be reasonable limits on the anti-discrimination right so long as they are “prescribed by law” and “demonstrably justified in a free and democratic society.”

The issue of whether sound rationing practices, taking into account prohibited grounds such as age and disability to make individualized assessments of need and ability to benefit, amounts to unlawful discrimination remains untested in New Zealand. A possible argument is that it is not discriminatory to treat people in different situations differently. But if resort were to § 5, the District Health Board or doctor would bear the onus of pointing to an objective and rational justification for the discrimination. In the context of rationing health services, the argument is that the equitable distribution of scarce resources would seem to be a legitimate objective. Giving priority to treatment to those patients who will gain most benefit would seem to be a justifiable means to that end, provided that the discriminatory criteria are rationally based and not disproportionate. The onus would be on providers to show an evidential basis in the medical literature for taking the discriminatory factor (age or disability) into account, and that the class of patients who share the prohibited criteria are not disproportionately excluded by reference to the evidence. The need for rationality and proportionality in terms of § 5 is key. The argument would be that the application of sound rationing criteria, which take into account prohibited grounds, would not offend the Bill of Rights standard, provided that there is an objective, evidential basis to justify taking the prohibited ground into account to that extent in the policy or guideline, or clinical decision.

Conclusion
The New Zealand legislature has been extremely reti-
cent in legislating guarantees of access to health services. Rather, the tendency has been for the fact of finite resources and the need to ration access to be explicitly acknowledged in legislation, notably in establishing the framework for public funding of health care and in the Code of Patients’ Rights. We have described the rise in interest since the 1990s in explicit rationing and a range of developments, including initial efforts to define a core of publicly available services, the development and refinement of priority-setting principles to underpin decisions, the growth in use of access guidelines, the development of a booking system for elective procedures, and moves towards a more systematic process within local health boards for adopting new health interventions.

We have assessed the two significant challenges to rationing decisions in the health sector. The South Auckland case did not proceed to litigation because the health authority’s resolve crumbled and treatment was provided. In the second case, Shortland, the hospital stood its ground, and the denial of dialysis treatment (to a man who subsequently died) was upheld in judicial review proceedings. The Court accepted the application of an access guideline as a defensible clinical judgement based on the patient’s best interests. Yet the decision masked the fact that the decision had been made in large part on resource grounds.

Since the reason for declining to provide dialysis was that the patient suffered from moderate dementia, as a result of which he was unable to administer treatment to himself, there was at least a prima facie case of disability discrimination to answer in Shortland. The ultimate issue was whether the guidelines themselves were discriminatory. It may be that the refusal to provide dialysis would have been justified by the “reasonable accommodation” defense. But whenever the criterion for allocation is an assessment of potential “benefit” in life-years, there is the potential for indirect discrimination based on age, disability, or other prohibited grounds such as ethnic origin. The result is that disproportionate numbers of older patients, patients with co-morbidities, or patients from ethnic minorities with poorer health status will fail to qualify for treatment. Whether sound rationing practices, taking into account prohibited grounds such as age and disability to make individualized assessments of need and ability to benefit, amount to unlawful discrimination remains untested in New Zealand. The issue will now fall to be determined under New Zealand’s Bill of Rights Act. Our contention is that the application of sound rationing criteria, which take into account prohibited grounds such as age or disability, would not offend the Bill of Rights standard, provided that there is an objective, evidential basis to justify the use of the differential factor in the policy or guideline, or clinical decision. The future sustainability of New Zealand’s public health system may depend to a significant extent on such a conclusion.

Acknowledgments

Grateful thanks to Paul Rishworth for his helpful comments on an earlier draft.

References


5. The balance of 22% of health expenditure in 2002 was funded from private health insurance (5.7%), out-of-pocket payments (36.1%), and not-for-profit organizations (0.3%). The percentage of health expenditure funded from private health insurance has risen markedly from 1.1% in 1979/80, see P. Davis and T. Ashton, Health and Public Policy in New Zealand (Auckland: Oxford University Press, 2001): at 10, to 5.7% in 2001/02, and has led to a two-tier system where insured, generally wealthier patients can usually access elective surgical procedures more quickly in private hospitals rather than wait for (publicly funded) public hospital treatment. The percentage of the population covered by private health insurance has been slowly declining, from an estimated 41% in 1994/95, id., to 33% in 2003/04, Ministry of Health, Health Report: Southern Cross Publication on Rebate for Private Health Insurance (Wellington: Ministry of Health, 2004): at 2.

6. “Treatment injury” includes personal injury suffered by a person seeking or receiving treatment from a registered health professional, which is caused by treatment but which is not a necessary part, or ordinary consequence, of the treatment, taking into account all the circumstances of the treatment, including (i) the person’s underlying health condition at the time of the treatment; and (ii) the clinical knowledge at the time of the treatment. See Injury Prevention, Rehabilitation, and Compensation Act 2001 (NZ) at § 32(1).


8. The terms, and others such as “limit-setting,” are used interchangeably in this paper.

9. See the New Zealand Public Health and Disability Act 2000 at § 3(2). The Minister of Health has power by written notice to require a District Health Board (DHB) to fund specified services, but any such notice may not require the supply of services to any named individual; see id. at § 33(1)(a).

10. See the Injury Prevention, Rehabilitation, and Compensation Act 2001 (NZ) at § 33(1)(d) and § 32(1)(c).

11. See the Code of Health and Disability Services Consumers’ Rights (NZ), clause 3, which provides: “(1) A provider is not in breach of this Code if the provider has taken reasonable actions in the circumstances to give effect to the rights, and comply with the duties, in this Code. (2) The onus is on the provider to prove that it took reasonable actions. (3) For the purposes of this clause, ‘the circumstances’ means all the relevant circumstances, including the consumer’s clinical circumstances and the provider’s resource constraints.”
12. See the Health and Disability Commissioner Act 1994 (NZ) at § 20(1)(f).
16. “Implicit” rationing occurs where care or provision of services is limited, but no explicit admission of limitation is made or justification for the limit is clearly expressed. The need for limit-setting is addressed on an informal basis – there are no explicit or formal criteria for making allocation decisions. The most notorious mechanism for implicit rationing is the waiting list. “Explicit” rationing, by contrast, is where an explicit or formal framework or allocation mechanism is employed, based on explicit criteria, with the aim of making consistent, fair and transparent decisions about access to treatment.
17. For this description of implicit rationing, see C. Ham and A. Coulter, “Where Are We Now?” in A. Coulter and C. Ham, eds., supra note 15, 239-50, at 249.
18. As is well known, an attempt at managed competition had recently been introduced in Britain, the core of which was the internal market or quasi-market and the separation of purchaser and provider roles. See W. Ranade, “Reforming the British National Health Service: All Change, No Change?” in W. Ranade, ed., Markets and Health Care: A Comparative Analysis (New York: Addison, Wesley, Longman, 1998): 101-21. Market mechanisms and incentives were introduced or strengthened in the health care systems of a group of advanced western states during the period, including Sweden, the Netherlands, and Germany; see id.
19. Public hospitals were renamed “Crown Health Enterprises” and charged with “operating as a successful and efficient business” and being “as successful and efficient as comparable businesses that are not owned by the Crown.” See the Health and Disability Services Act 1993 (NZ) at §§ 11(1) and (2)(d).
21. See the Health and Disability Services Act 1993 (NZ) at § 4(a).
22. Id.
23. The Committee’s remit was to advise the Minister of Health on “the kinds, and relative priorities, of personal health services and disability services that should, in the committee’s opinion, be publicly funded.” See Health and Disability Services Act 1993 (NZ) at § 6(1a). In 1996 advice on priorities for public health services was included within its brief. The current national advisory committee on health and disability, established under the New Zealand Public Health and Disability Act 2000, § 13, retains this advisory function in relation to “the kinds, and relative priorities” of services that should be publicly funded.
24. See Third Report of the National Advisory Committee on Core Health and Disability Support Services, Core Services 1995/96 (National Advisory Committee on Health and Disability, August 24, 1994). Reflecting the different task, the Committee underwent a name change about this time, thereafter becoming the National Health Committee.
25. These principles were set out in a consultation document widely circulated in 1993, The Best of Health 2 (National Advisory Committee on Health and Disability, 1993) and were adopted by the Committee and advised to the Minister of Health. See Third Report of the National Advisory Committee on Core Health and Disability Support Services, supra note 24, at 8.
28. The Committee commissioned a report on waiting lists from three clinicians, which proposed the introduction of booking to give some certainty to patients and promoted prioritization according to need and ability to benefit. See G. Fraser, P. Alley, and R. Morris, Waiting Lists and Waiting Times: Their Nature and Management (Report to the National Advisory Committee on Core Health and Disability Support Services, 1993).
30. The booking system was reviewed independently from an ethical perspective, and that review was itself reviewed. Both reports concluded that despite unresolved questions and ethical dilemmas (particularly the tension between need and ability to benefit), the project was a clear improvement on the old system of implicit rationing by waiting list, which lacked transparency and accountability, and the potential to abuse. The new system was more transparent and equitable and, by publicly providing information on levels of unmet need, provided more opportunities for advocates and the public to participate in open discourse on ways to improve the delivery of health care. See D. Evans and N. Price, The Ethical Dimensions of the National Waiting Time Project (A Report for the Health Funding Authority, 1999); I. Kawachi, Review and Commentary on the Ethical Dimensions of the National Waiting Time Project (A Report for the Health Funding Authority, 1999).
32. Pharmac is governed by the New Zealand Public Health and Disability Act 2000, §§ 46-53.
33. In the year ended June 31, 2004 its expenditure on pharmacueticals was $NZ 534 million; see Annual Report of Pharmaceutical Management Agency for Year ended 31 June 2004. For a more detailed description, its activities, and success, see Bloomfield, supra note 15, at 32-35.
35. Bloomfield comments that the main weakness of the process was that the underlying community principles had not been submitted to public consultation. See Bloomfield, supra note 15, at 23.
36. Maori are New Zealand’s indigenous people, to whom the Crown owes special duties under a national founding document, the Treaty of Waitangi. The fifth principle, the need to improve the health status of Maori, reflected their persistently lower health status compared to non-Maori New Zealanders, apparent in disparities in life expectancy and the prevalence of cardiovascular disease, cancer, diabetes, and chronic lung diseases; see Ministry of Health, Decades of Disparity: Ethnic Mortality Trends in New Zealand 1980-1999 (Wellington: Ministry of Health, 2003). One objective of the New Zealand Public Health and Disability Act 2000 is to reduce health disparities by improving the health outcomes of Maori and other population groups, see § 3(1)(b), although § 3(3) qualifies this by stating that: “For the avoidance of any doubt, nothing in this Act-(a) entitles a person to preferential access to services on the basis of race.”
Tenbensel, supra note 15, at 195.

38. The reforms were put in place by the New Zealand Public Health and Disability Act 2000.

39. See New Zealand Public Health and Disability Act 2000 at § 3(1)(c)(i). District Health Boards are at the same time required to be fiscally prudent (§ 41), attempt to stay within their budgets (§§ 41, 42), and observe and implement government priorities for the sector (§§ 38(6), 39(8)). These dual accountabilities of Board members to the government and to local communities have caused some tension, and raise an issue felt particularly by elected members of Boards, as noted in research undertaken to chart the progress of and evaluate the reforms as they are implemented. See Health Reforms 2001 Research Team, Interim Report on Health Reforms 2001 Research Project (Health Reforms 2001 Research Team, November 2003); at 36-37.

40. It is a function of DHBs to regularly assess and monitor the health status of their resident populations and their needs for services; see New Zealand Public Health and Disability Act 2000, s 230(1)(g). Health needs assessments are to be used to inform DHBs’ District Strategic Plans (DSP). See § 38(3)(a). The DSP states the objectives of the DHB for the five- to ten-year period, and is to be made publicly available, as are any amendments. These plans are to be reviewed at least once every three years. See § 38.

41. See Tenbensel, supra note 15, at 204.


43. The majority of DHBs are in deficit, which they are required to progressively reduce. In this environment, prioritization is related to disinvestment decisions rather than to investing in new services.

44. See Health Reforms 2001 Research Team, supra note 39, at 48-51. The Report provides evidence for the 2001-03 period only.

45. See Health Reforms 2001 Research Team, supra note 39, at 49. See also Ashton, et al., supra note 42, at 28.

46. For example, the introduction of a national hepatitis screening program, the breast screening program, and the acellular pertussis vaccine, which were the result of decisions made by government. For a summary of how decisions about new health technologies have been made in New Zealand, see Bloomfield, supra note 15, at 25-32, and at 37-41.


48. Factors include consumer expectations and lobbying, adoption by other DHBs, “technology creep,” whereby existing interventions are replaced by newer modified and usually more expensive versions, expectations of new clinical staff, adoption by the private sector. See id. at 15-16.

49. The Committee found that decisions are often based on one or two articles from reputable journals or clinical trials, rather than a synthesis of all available evidence. See id. at 3.

50. Id. at 24.

51. Id. at 20-31.

52. The Committee identified the key attributes for robust decision-making processes, based on Daniels and Sabin’s “accountability for reasonableness.” See N. Daniels and J. Sabin, “The Ethics of Accountability in Managed Care Reform,” Health Affairs 17, no. 5 (1998): 50-64. It recommended development of a resource for hospital decision-makers on how to establish robust processes consistent with these attributes. See the Committee’s report, Decision-Making about New Health Interventions, supra note 47, at 26-27.

53. The suggested options included: a web-based searchable library to act as a repository of relevant international and national generated evidence and health technology assessment information; and a brokerage agency to provide a rapid source of such information for instances in which the web-based library was inadequate. See id. at 33-35.


55. An interesting example was the recommendation by the Chief Ombudsman that a regional health authority not resale from a commitment to fund a sex-change operation for a transsexual. The authority had revoked its decision to fund the surgery after reclassifying the procedure. The Ombudsman considered the decision to reclassify the procedure “unreasonable.” See File No W 40304 (November 2000).


57. Removal of waste products through a fluid exchange system attached to the abdomen. Cf. haemodialysis, which is removal of waste products by circulating the blood through a filter system, traditionally undertaken as an in-hospital procedure.

58. The health team consisted of a renal physician, a social worker, a renal nurse, a Maori liaison officer, an occupational therapist, and a psychologist. A second renal physician provided a second opinion.

59. See Northern Regional Health Authority’s Guidelines for Entry into Northern Region’s End Stage Renal Failure Programme (Northern Regional Health Authority, 1996). The guidelines were a twelve-page document developed in 1996 by a committee established by the regional funder. Members consisted of doctors, health authority representatives, medical ethicists, lawyers, patient group representatives, nurses, and iwi (Maori tribal members). Co-author of this paper, Ron Paterson, was a member of the committee. The committee met monthly over an eighteen-month period and reviewed issues such as medical resources, patient best interests, medical practice, medical ethics, medical law, and human rights. Drafts were widely disseminated and the Human Rights Commission was kept fully informed. Justice Salmon concluded that the guidelines were the subject of thorough and appropriate study before they were adopted; see Shortland v. Northland Health Ltd (No. 2) (unrep, HC Whangarei, M No. 75/97, 6 November 1997, Salmon J.), at 10.

60. The guidelines, indicating legal input in this respect, stated that they were “guiding principles” only, and provided for “exceptional circumstances” in which a “positive decision” could be made “to offer treatment even though it is unlikely to have benefit greater than two years.” See Shortland v. Northland Health Ltd, supra note 27, at 438.


62. The evidence was that attempts were made to teach Williams to perform CAPD over several weeks, but he was unable to learn or retain the basic concepts. He disconnected his CAPD on two occasions, on one occasion causing life-threatening peritonitis; see Shortland v. Northland Health Ltd, supra note 27, at 438; and Shortland v. Northland Health Ltd (No. 1), supra note 61, at 5.

63. The evidence indicated, however, that it had proved impossible to ensure a family member took responsibility for his care during the assessment period when efforts had been made to train Williams in the procedure; see Shortland v. Northland Health Ltd (No. 1), supra note 61, at 5.


65. See Shortland v. Northland Health Ltd (No. 1), supra note 61; and Shortland v. Northland Health Ltd (No. 2), supra note 59. Like the Court of Appeal, Justice Salmon announced his decision with a short statement of reasons (on October 10, 1997), and provided a full statement of reasons later.


67. Resumption of dialysis, had the Court of Appeal ordered it, would in all probability have been too late to save his life in any event, although the legality of its withdrawal in the first place was always in issue. The Health and Disability Commissioner subsequently found fault with the process but not the outcome. In her opinion, Northland Health had failed to provide services that took into ac-
count Williams’ cultural and spiritual needs, and to comply with relevant standards for consultation with family, in breach of Rights 1(3) and 4(2) of the Code of Patients’ Rights (Case 97HD C8872, June 28, 1999), available through <www.hdc.org.nz/opinions>.

68. Shortland v. Northland Health Ltd (No. 1), supra note 61, at 9-10, citing the Health and Disability Services Act 1993 (NZ), §§ 4(a) and 8(3).


70. See Re J (A Minor) (Child in Care: Medical Treatment) [1992] 3 WLR 507 (CA).

71. See id., at 517 (L. J. Donaldson) and 519 (L. J. Balcombe); R v. Cambridge District Health Authority ex parte B [1991] 1 WLR 899 (CA), at 906.


75. Id. at 14 (emphasis added).


77. The New Zealand Bill of Rights Act 1990, § 8, provides: “Right not to be deprived of life – No one shall be deprived of life except on such grounds as are established by law and are consistent with the principles of fundamental justice.”

78. Shortland v. Northland Health Ltd (No. 2), supra note 59, at 17. Causation theories such as this are of course controversial. See Auckland Area Health Board v. Attorney-General [1993] 1 NZLR 235 (HC), at 248; Airedale NHS Trust v. Bland [1993] All ER 821 (HL), at 893 (Lord Mustill). Alternatively, the High Court held the decision was in accordance with fundamental justice, because it was made clear to Williams when he was placed on dialysis initially that the purpose was for assessment only; hence no legitimate expectation arose. And there was no breach of the principles of natural justice as the family had ample opportunity, which they exercised, to attempt to persuade the hospital to keep him on dialysis; see id. at 20-21.

79. The Crimes Act 1961 (NZ), § 151(1), states that “[e]very one who has charge of any other person unable, by reason of...sickness...to withdraw himself from such charge, and unable to provide himself with the necessaries of life, is...under a legal duty to supply that person with the necessaries of life, and is criminally responsible for omitting without lawful excuse to perform such duty if the death of that person is caused, or if his life is endangered or if his health permanently injured, by such omission.”

80. [1993] 1 NZLR 235 (HC).

81. It quoted the aim of the guidelines “to ensure that, so far as possible within the available resources, all patients are offered access to the treatment modality which is most suitable clinically and socially and which offers the greatest opportunity to benefit”; see Shortland v. Northland Health Ltd (No. 2), supra note 59, at 437 (emphasis added).

82. Id. at 442-43.

83. Id. at 439.


85. Id. at 14, at 442.

86. Right 7(2) of the Code provides: “Every consumer must be presumed competent to make an informed choice and give informed consent, unless there are reasonable grounds for believing that the consumer is not competent.”

87. Right 7(4) of the Code provides: “Where a consumer is not competent to make an informed choice and give informed consent, and no person entitled to consent on behalf of the consumer is available, the provider may provide services where – (a) It is in the best interests of the consumer; and (b) Reasonable steps have been taken to ascertain the views of the consumer; and (c) Either, – (i) If the consumer’s views have been ascertained, and having regard to those views, the provider believes, on reasonable grounds, that the provision of the services is consistent with the informed choice the consumer would make if he or she were competent; or

(ii) If the consumer’s views have not been ascertained, the provider takes into account the views of other suitable persons who are interested in the welfare of the consumer and available to advise the provider.”

88. A codification of the principle in F. v. West Berkshire Health Authority [1990] 2 AC 1 (HL), at 75-76 (Lord Goff).

89. Had the fact of making a choice between patients because of scarce resources been acknowledged, the issue would then have been whether breaches of Rights 7(5) and (4) were excused by the Code’s clause 3; see supra note 11.

90. See the line of authority for heightened judicial scrutiny (“the super-Wednesbury approach”) when fundamental rights are engaged, especially the right to life: R v. Secretary of State for the Home Department, ex parte Brind [1991] 1 AC 696 (HL), at 745-49; R v. Ministry of Defence, ex parte Smith [1996] 1 All ER 257, at 263; R. (Wilkinson) v. Responsible Medical Officer, Broadmoor Hospital (2001) 65 BMLR 15 (CA).

91. See Airedale NHS Trust v. Bland [1993] 1 All ER 821 (HL), at 871 and at 883, referring to Bolam v. Friern Hospital Management Committee [1957] 2 All ER 118.


93. See Re A (Medical Treatment: Male Sterilization) [2000] 1 FLR 549 (CA), at 555 and In re S (Adult Patient: Sterilization) [2000] 3 WLR 1288 (CA), at 1296.

94. Supra note 27, at 443.


96. Except where there is a legal guardian or power of attorney available to give consent to treatment of the patient.

97. Supra note 27, at 443.

98. The “definitional approach” is to define the relevant right narrowly, to avoid its being engaged on the facts of a case or to avoid conflict with other rights, rather than to resolve such conflicts by reference to the excepting grounds within the right itself (“the principles of fundamental justice”) or by reference to the “justified limitations” on rights provision in s 5 of the New Zealand Bill of Rights Act 1990. The latter is referred to as ad hoc balancing. See Re J (An Infant) [1996] 2 NZLR 134 (CA), at 145-46.

99. Supra note 27, at 445.

100. Id. at 445.

101. The qualification being that deprivation of life is consistent with “the principles of fundamental justice”; see the New Zealand Bill of Rights Act 1990, § 8, supra note 77.

102. See the New Zealand Bill of Rights Act 1990 at § 5, which provides: “Justified limitations – Subject to section 4 of this Bill of Rights, the rights and freedoms contained in this Bill of Rights may be subject only to such reasonable limits prescribed by law as can be demonstrably justified in a free and democratic society.” Section 5 was adopted from the Canadian Charter of Rights and Freedoms at § 1.


105. To borrow Syrett’s description of implicit rationing, see id. at 293.

106. Daniels and Sabin’s four conditions for “accountability for reasonableness” are: (1) Publicity, which requires decisions regarding coverage for new technologies (and other limit-setting decisions) and their rationales to be publicly accessible; (2) Reasonableness, requiring that the rationales for coverage deci-
sions should rest on evidence, reasons, and principles that all fair-minded parties can agree are relevant to meeting people's needs fairly under resource constraints; (3) Appeals, which requires a mechanism for challenge and dispute resolution and an opportunity for revising decisions; and (4) Enforcement, which requires voluntary or public regulation of the process to ensure that conditions 1-3 are met. See Daniels and Sabin, supra note 52; N. Daniels, “Accountability for Reasonableness in Public and Private Health Insurance,” in A. Coulter and C. Ham, eds., The Global Challenge of Health Care Rationing, supra note 15, at 92-93.

107. For a further example of a decision, in which the Court of Appeal found a governmental action lawful while not explicitly acknowledging the issues to relate, at least in part, to a decision allocating limited resources, see Daniels v. Attorney-General (2003) 3 NZLR 742 (CA). In the education rather than the health field, the decision differed from Shortland in that it involved a statutory entitlement to “free enrolment and education,” which the Education Act 1989 (NZ) extended to children with special educational needs. Rejecting the interpretation of the court below, the Court of Appeal held that the statutory right did not create a free-standing, general right to “regular, not clearly unsuitable and systematic education.” The statutory right was not justiciable in terms of the suitability of the education provided for particular children. This interpretation rendered the statutory right virtually meaningless, for the content of the right that the Court was prepared to acknowledge added nothing to other specific duties in the Act. Underlying the decision is an apparent unwillingness to become involved in reviewing the relative allocation of funding as between students with differing levels of disability and special needs in the context of a policy that had significantly increased funding for special education overall.


109. The International Covenant on Civil and Political Rights, ratified in New Zealand in 1978, was a major influence on the New Zealand Bill of Rights Act 1990.

110. The New Zealand Bill of Rights Act 1990 applies to the three branches of government – executive, legislative, and judicial – and also to bodies exercising a public function (see § 3). Like the Human Rights Act 1989 (UK), the Act is not supreme law. A court cannot invalidate a statute on the ground of inconsistency with the New Zealand Bill of Rights (see § 4), although those subjects to the Act, including courts, are instructed to read a statute consistently with the guaranteed rights (see § 6).

111. See the Human Rights Act 1993 (NZ) at § 44(1).

112. Id. at § 21(1).

113. Id. at § 21(1)(b)(v).

114. The family had laid a complaint of disability discrimination prior to bringing the judicial review proceedings, which had been withdrawn after the Human Rights Commission undertook a process of conciliation between the parties.

115. See the relevant section of the guideline, which is quoted in the journal of Medical Ethics 31, no. 7 (2005): 373.

116. Such as Mr Soobramoney, who had heart and vascular disease, which meant that his ability to benefit was compromised, and he fell outside the guidelines. See Soobramoney v. Minister of Health, KwaZulu-Natal (1998) 1 SA 765 (CC).


118. There was a pattern of increasing life-expectancy gaps for both Maori and Pacific people compared to non-Maori, non-Pacific people between 1981 and 1999. For Maori, the male life-expectancy gap in 1999 was 9.5 years; for females, 9.8 years. Chronic diseases contribute the major share of the disparity in life expectancy between Maori and non-Maori, non-Pacific people. See Ministry of Health, Decades of Disparity: Ethnic Mortality Trends in New Zealand 1980-1999 (Wellington: Ministry of Health, 2003): at 14, 22, and 42.

119. See the relevant section of the guideline, which is quoted in the journal of Medical Ethics 31, no. 7 (2005): 373.

120. Employment status is defined as meaning: “(i) Being unemployed; or (ii) Being a recipient of a benefit or compensation under the Social Security Act 1964 or the [Injury Prevention, Rehabilitation, and Compensation Act 2001]” and family status is defined as meaning: “(i) Having the responsibility for part-time care or full-time care of children or other dependants; or (ii) Having no responsibility for the care of children or other dependants; or (iii) Being married to, or being in a relationship in the nature of marriage with, a particular person; or (iv) Being a relative of a particular person.”

121. New Zealand’s National Health Committee concluded after consultation that social factors have a place, but a limited one, in deciding priorities. A 10% weighting in priority scoring was allocated in its Priority Criteria Projects in the 1990s to ability to work, care of dependants, and living independently; see W. Edgar, “Rationing Health Care in New Zealand – How the Public has a Say,” in A. Coulter and C. Ham, The Global Challenge of Health Care Rationing, supra note 15, at 183-84 and 191-92, supra note 15.

122. For example, arguably the most influential modern theory of justice considered that distribution of “social values” should be equal unless an unequal distribution would favour the least well off; see J. Rawls, A Theory of Justice (Cambridge, MA: Harvard University Press, 1971).


125. See the Human Rights Act 1993 (NZ) at § 65.

126. “Genuine justification” is a defense to § 44 (direct discrimination), but is available on a case-by-case basis only, not as a blanket ruling. See the Human Rights Act 1993 (NZ) at § 57; “Good reason” is available as a defense to indirect discrimination. See id. at § 65.

127. See the Human Rights Amendment Act 2001 (NZ). The Amendment was necessary because the government’s exemption from certain grounds of prohibited discrimination in the Human Rights Act was due to expire, and the whole range of governmental activity was to become subject to the Act. This would result in serious problems, for example in respect of social assistance payments, which take into account prohibited grounds. Note that public hospitals and their employees were not included within that exemption, although funding authorities and the Ministry of Health were.

128. See the law reform report leading to passage of the Human Rights Amendment Act 2003 (NZ): P. Cooper, P. Hunt, J.
Joanna Manning

129. Id. at paragraphs 27 and 73.
130. In two cases in 2003, Auckland City Hospital denied dialysis treatment to two foreign patients who were lawfully in New Zealand on temporary entry permits. It applied a government policy to restrict such treatment to New Zealand citizens and permanent residents. Pursuant to the policy, foreign patients who presented would be stabilized but not offered ongoing dialysis. Both patients were treated differently from New Zealand citizens by reason of a prohibited ground (ethnic or national origin), which appears *prima facie* discriminatory. The issue is whether it is a reasonable limitation on their anti-discrimination right for the government to decide as a matter of policy to spend its finite health dollars on New Zealand patients. Amid intense media interest, both cases were resolved without recourse to litigation. In the first case, the Minister of Immigration apparently succumbed to pressure and issued the patient with the necessary permit, so that he qualified for public health services; see “Man’s Life in Balance as Health Policy Debated,” New Zealand Herald, May 10, 2003; “Tuvaluan Man Given Temporary Permit to Continue Life Saving Treatment,” New Zealand Herald, May 12, 2003. The second case involved an eighteen-year-old Fijian-born Indian girl, who came to New Zealand to receive a kidney transplant from her aunt. The latter unfortunately suffered a stroke soon after arrival and was no longer a suitable donor. As the girl’s insurance cover ran out, Auckland City Hospital advised of its intention to cease dialysis treatment. After media publicity, a private benefactor agreed to pay the costs of treatment for twelve months until a suitable donor could be found; see “Kidney Girl in Fight for Insurance,” New Zealand Herald, July 15, 2003; “Kidney Girl’s Time Up,” New Zealand Herald, July 30, 2003; and “A Dream Comes True For Kidney Patient,” New Zealand Herald, July 11, 2003.

131. See Daniels v. Attorney-General (No. 1) (unrep, HC Auckland, M No. 1615-SW99, 3 April 2002, J. Baragwanath), at paragraph 97, in which the High Court took this view.
132. It is beyond the scope of this paper to discuss in detail the question of whether the New Zealand Bill of Rights Act 1990 applies to the actions of health professionals in public hospitals administering publicly funded health services. It is generally considered, however, that they perform a public function in terms of § 3(b) of the Act, and hence that the Act applies to them; see P. Rishworth, G. Huscroft, S. Optican, and R. Mahoney, *The New Zealand Bill of Rights* (Melbourne: Oxford University Press, 2003): at 91.

134. The New Zealand courts approach this issue in very much the same way as do Canadian courts under § 1 of the Canadian Charter of Rights and Freedoms, and have adopted the *Oakes* test. See *R v. Oakes* [1986] 1 SCR 103 (SCC); *Moonen v. Film and Literature Board of Review* [2000] 2 NZLR 9 (CA).
Access to Health Care in the Netherlands: The Influence of (European) Treaty Law

André den Exter

1. In the Netherlands, access to healthcare has been guaranteed by social health insurance legislation. But since the introduction of the Health Insurance Act ("Ziekenfondswet") in the 1960s, the health insurance system has been in a state of flux. Numerous reforms have changed the system gradually, of which the latest is the introduction of a competitive health insurance scheme for the entire population.

Cutting across the various reforms has, however, been the goal of access to healthcare services as defined by international treaty law, including European Union law. In particular the leverage of Community law in strengthening the patient’s right to healthcare is remarkable. Since the European Court of Justice (ECJ) has accepted that healthcare should be considered as a service in terms of the EC Treaty, rights to healthcare have become inextricably linked with the free movement principles and are no longer restricted to the jurisdiction of the country of origin.

Hereafter, this article examines the consequences of international and European Union law for claiming access to healthcare in the Netherlands.


The Dutch healthcare system is based on the principle of solidarity. Solidarity means that all members of society in need must have access to healthcare, regardless of their ability to pay. Solidarity is not a woolly notion about the common good. It has a specific meaning that a healthcare system is organised and managed on the basis of universal access, without risk selection, based on income related premiums and with no significant differences in the benefit package. It is a concept enshrined in Dutch statutory health insurance law and recognised by the decisions of the judiciary, including the European Court of Justice, as described below.

Due to the prohibition of constitutional review by the judiciary, the Dutch legal debate on access to healthcare starts from a statutory law perspective. More specifically, that debate primarily focuses on health insurance issues (access to the health insurance scheme, the scope and nature of the benefit package, contribution rates, etc.). That does not mean that the constitutional health provision is of no importance, but in terms of healthcare claims statutory law and international legal standards are decisive.

Generally, the principle of access to health care has been interpreted as a social right, which means that the

André den Exter, J.D., is a lecturer in European law, Radboud University Nijmegen, the Netherlands, Correspondence: a.denexter@jur.ru.nl
competent authorities in the Netherlands have quite a broad discretion in determining how and when this social right will be implemented. Based on both the solidarity principle and the equality principle, equal access to healthcare services has been enshrined in social insurance law (e.g., the Health Insurance Act, “Ziekenfondswet”) and formulated into healthcare entitlements as defined by (secondary) law. The “Ziekenfondswet” establishes a statutory insurance scheme covering the insured for curative care. Eligibility for coverage has been determined by a statutory income ceiling (set at €32,000 for 2005), which applies to more than sixty percent of the population. Besides the statutory nature, the law defines the nature of care in terms of entitlements (e.g., general practitioner care, paramedical care, specialist medical care, obstetric care, etc.), whereas the details are regulated by the Health Insurance (Treatment and Services) Decree and associated ministerial regulations. These forms of care have been further developed in the health insurance policies of individual insurance funds/companies. The devil has been in the details; these policies are crucial in litigation procedures.

In principle, the insured may claim the types of care as benefit-in-kind entitlements provided by contracted health providers. The insured have free choice of provider, but that is restricted to contracted providers only. However, in exceptional cases the insured may opt for a non-contracted provider, for instance in the face of long waiting times, and will receive a cash benefit (reimbursement model). Due to jurisprudence of the European Court of Justice, however, the reimbursement rules had to be amended (this is discussed further below). One of the reasons is the statutory obligation of the insurance fund to guarantee access to medical care under the insurance scheme. This obligation forms the essence of the benefit-in-kind healthcare system, for which the insurer is accountable and – in case of non compliance – may be held liable. It is therefore a principle of sound management that the benefit-in-kind insurer purchases sufficient care in and/or outside its area of activity. Except for force majeure, the insurance fund cannot be released from its obligation to provide for access to healthcare. The insurance fund cannot defend itself on the basis that it had insufficiently contracted for care or allocated insufficient funds to a hospital facility. It can, however, claim force majeure in the case of restrictive residential capacity regulation set by the government as this is viewed as an externality factor beyond the powers of insurers to affect. Nonetheless, given the possibility of contracting health care abroad, such an appeal will be likely be less successful in the future.

International (Social Security) Law

Apart from national health insurance law, the legal framework on social health insurance has been highly influenced by international social security law, as stipulated in several treaties and conventions. These international documents conceptualise the measures that should be undertaken by Member States in order to achieve the full realisation of the right to healthcare.

Generally, the principle of access to health care has been interpreted as a social right, which means that the competent authorities in the Netherlands have quite a broad discretion in determining how and when this social right will be implemented.

The formulation used include a certain responsibility upon states with regards to the allocation, finance and provision of health services, but these provisions are generally interpreted as a non-enforceable legal entitlement to healthcare. Therefore, the noble aspirations of universal access to healthcare must be interpreted within the context of the treaties, meaning a basic level of healthcare, interpreted in accordance with the formulation, ratio and implementation of legal norms and economic capacities of a given society. In truth, the domestic economic capacities and the non-enforceability of these international treaty provisions temper the actual meaning of the right to healthcare and thus access to healthcare services.

In the Netherlands, however, the importance of international treaties on social rights has gradually increased since the concept of “direct effect” or the notion of self-executing treaty provisions. On several occasions, the Dutch judiciary accepted that international treaty law on social rights may have binding effect on all persons (“erga omnes” provisions) and certain incorporated rights (on labour issues and social security benefits) have been subject to judicial review. Particularly the direct effect of norm setting treaties, such as ILO conventions on social security has raised some controversy, both in the legal doctrine, as well as at the courts. With respect to health, the highest administrative court, the Central Appeals Tribunal (“Centrale Raad van Beroep”) ruled that the ILO Convention on social security minimum standards and maternity care
include self-executing treaty provisions that can create direct rights to individuals, enforceable by the Court. Decisive criteria whether norm setting treaties or treaty provisions are self-executing are their (instructive or imperative) nature, and the (concrete) wording of the specific provision. Therefore, the reliance on the self-executing effect of ILO social security treaties provide Dutch citizens with a limited claim to enforce the social right to healthcare before national courts. Conversely, the judiciary rejected such reliance repeatedly in case of the International Convention on Economic, Social and Cultural Rights (ICESCR), since the treaty provision are insufficiently precise, and the instructive nature provides States with a broad margin of appreciation to fill in the necessary steps in order to realise these rights. So far, the judiciary has continued that line of reasoning and is not willing to incorporate the concept of “progressive realisation” of social rights, as suggested by the treaty Committee in its General Comments and Concluding Observations (on health).

Alternatively, citizens may also rely on the European Human Rights Treaty, the European Convention on Human Rights (ECHR) when claiming access to and/or the reimbursement of a certain medical intervention or medicines. On several occasions, particularly when the medical intervention or medicine was not included in the social health insurance benefit package, the ECHR treaty provisions have been invoked in order to enforce a healthcare claim on grounds of the private or family life provision (Article 8). For instance, in Sentges v. The Netherlands, a teenage boy with multiple handicaps, Nicki Sentges, complained that his request for a robotic arm was denied. He submitted that under Article 8 the authorities were under a positive obligation to provide him with this medical device, arguing that the concept of private life, as interpreted by the Court, encompassed notions pertaining to the quality of life, including personal autonomy, and the right to establish and develop relationships with other human beings. Sentges argued that the constraints on him were unacceptable as he was never able to be alone and his total dependency on others “forced him to establish and develop friendships that he might not chosen had he not been disabled.” While the essential object of Article 8 is to protect the individual against arbitrary interference the Court has held that this provision may also include positive obligations inherent in effective respect for private or family life. These obligations may involve the adoption of measures designed to secure respect of private life. But in order for there to be a positive obligation on the state there needs to be a “direct and immediate link” between the measures sought by the applicant and the latter’s private life. Regrettably, the Court declined to decide whether such a link had been established (although it noted that the national Central Appeals Tribunal had accepted such a link). Instead, the Court concluded that particularly in issues that involve the assessment of priorities of limited healthcare resources, national authorities enjoy a wide margin of appreciation since they “are in a better position to carry out this assessment than an international court.” This is also called the “fair balance” test of the competing interests of the individual and the community as a whole, including the costs of healthcare. In the present case, the Court considered that the provision of a robotic arm fell within the margin of appreciation since the applicant has access to the standard package of health care provided by the Health Insurance Act, i.e. an electric wheelchair with an adapted joystick. Only in case of manifest unreasonable outcomes will the Court consider intervening. This could occur, for example, if an applicant was denied a life-saving medicine and it is possible for the cost thereof to be met by the State. In the case of life essential treatment or medicines, one may also consider an appeal to the right to life, Article 2 of the Convention.

Only in exceptional cases, the alternative approach provided by the Human Rights Court may be effective, but given the Court’s understandable hesitation to link the healthcare right with individual rights, *nominatim* life and private life, and its “fair balance” test, makes that it is extremely difficult, if not impossible, to enforce healthcare claims. Claiming health insurance entitlements based on statutory law provide a more successful remedy to realise one’s right to healthcare. Due to the influence of the European Court of Justice, we will see that these claims are not restricted to the Dutch territory.

### 3. European Union Law and Healthcare

#### 3.1. Public Health Law

The European Union (previously, the European Community) was established with the ratification of the Treaty of Maastricht (1992). Since Maastricht, the European Union (EU) Treaty has been modernised and replaced by the so-called Treaty of Amsterdam (1997), and the Treaty of Nice (2001). These Treaties have created a new, distinctive legal order that can be distinguished from traditional international law by its content, instruments and sources of law.

For historical reasons, the EU has no general competence to regulate in the field of health. The original Treaty of Rome establishing the European Communities did not foresee a “European” health system. Health policy, notably the issue of access to healthcare services and facilities has been regulated by the domestic legal orders of individual Member States. Community competences with respect to health were mainly based on general treaty provisions as far as they con-
cerned the functioning of the common market.\textsuperscript{21} It was only in the Treaty of Maastricht that the EU received supranational competence to run a public health policy of disease prevention and health promotion (Article 129 EC Treaty of Maastricht). The conferred public health tasks to the EU were restricted however by the principle of subsidiary, \textit{i.e.} only if and insofar as the objectives of the proposed action cannot be sufficiently achieved by the Member States.\textsuperscript{23} Thus, the Treaty constrains EU jurisdiction \textit{vis-à-vis} health.

The Amsterdam amendment (1997), introduced considerable changes although the lack of legal precision of Article 152(4)(a) still leaves scope for interpretation.\textsuperscript{24} Article 152 EC of the Treaty of Amsterdam replaced the Maastricht public health provision, Article 129. As a result, Union action will no longer be limited to measures or activities, which are preventive in nature.\textsuperscript{25} Since "such action shall cover the fight against the major health scourges, by promoting research...as well as health information and education" (section 1), the Union may also take positive steps in order to improve public health. Further extended Union public health actions include measures, \textit{inter alia}, setting standards of quality and safety of organs and substances of human origin, blood and blood derivatives and in the fields of veterinary and phyto sanitary health.\textsuperscript{26}

In the fifth section it is mentioned that Union action in the field of public health "shall fully respect the responsibility of the Member States for the organisation and delivery of health services and medical care." It re-emphasises the reluctance of Member States to hand over their competences in organising and financing their respective national healthcare system. Consequently, individual claims to access healthcare services cannot be based on Article 152.

3.2 The Free Movement Principles

Although national authorities have, in theory, exclusive competencies \textit{vis-à-vis} the organization, financing and delivery of health care services, this jurisdiction is affected by policy decisions made at the European level and by provisions of EU law designed to realise the internal market.

According to article 14(2) of the EC Treaty, the internal market "shall comprise an area without internal frontiers in which the free movement of goods, persons, services and capital is ensured in accordance with the provisions of this Treaty." Under certain circumstances, Union citizens may even derive specific rights emanated from the "free movement" provisions. For patients, the most relevant provisions are the free movement of persons, services and goods.\textsuperscript{27} The internal market provisions' impact on the health sector is, however, incomplete and differs by provision.

With regard to the free movement of persons, relevant Treaty provisions include the freedom of movement for "workers" (art. 39-42 EC) and the rights of "establishment" (art. 43-44 EC). These provisions have, in turn, been further substantiated by secondary legislation.\textsuperscript{28} Free movement for workers shall entail, \textit{inter alia}, the right to stay in a Member State for the purpose of employment in accordance with the provisions governing the employment of nationals of that State laid down by law, regulation or administrative action.

The free movement provisions are relevant not only to health professionals, but also to patients. Article 22 of co-ordination regulation 1408/71 (Regulation 1408/71) in conjunction with Article 39 EC, entitles cross-border workers to access the health care system in their country of residence; emergency care in case of temporary residence abroad; and pre-authorized care abroad by the patient’s insurer or the competent (national) health authority.\textsuperscript{29} Since free movement is not restricted to "workers," relatives, tourists and other categories of EU citizens can also make an appeal to benefit from this provision. Although Regulation 1408/71 aims to coordinate the different social security systems in the Member States, free movement of patients remains problematic. A major problem countries face with cross-border health care is how to regulate and finance this varied care. Some Member States fear an influx of patients from those Member States lacking facilities and/or providing lower-quality care. Rulings from the Court of Justice, simplifying patient mobility, have only strengthened this fear.

4. The ECJ’s Rulings on Patient Mobility and its Consequences for the Dutch Social Health Insurance System

4.1 Patient Mobility

The European Court of Justice (ECJ) is entrusted with, \textit{inter alia}, the provision of preliminary rulings.\textsuperscript{30} On references of the parties in the main proceedings before the national court, the ECJ, by means of a preliminary ruling, is exclusively competent to pronounce the interpretation of the Treaty provisions if a question on this subject is raised. As such, the Court is able to promote the uniformity of interpretation of Union law in the Member States.

In quite a number of cases, the Court’s jurisdiction has been invoked in national judicial procedures on patient mobility.\textsuperscript{31} In a typical case an insured complains that his/her insurance fund has imposed a restrictive condition or denied approval to receive care abroad and thus questions the legitimacy of the Regulation’s pre-authorization requirement in view of the fundamental freedoms. Since the precedent-setting case of Decker and Kohll, the ECJ has been confronted
with a growing number of cases questioning the conformity of the Dutch social health insurance system with the free movement principles. In the joined case of Smits and Peerbooms, the Dutch social health insurance fund refused Mrs. Smits reimbursement for specific multidisciplinary hospital treatment costs incurred in a German clinic, specialised in Parkinson’s disease. Justifying its decision, the social insurance fund said that satisfactory and adequate treatment for Parkinson’s disease was available in the Netherlands and that the specific clinical treatment provided in Germany provided no additional advantage. Thus, it was not “medically necessary” that Mrs. Smits undergo treatment at the German clinic.

The second claimant, Mr. Peerbooms fell into a coma following a road accident. He was transferred to a hospital in the Netherlands and then transferred in a vegetative state to the university clinic in Innsbruck in Austria where he received a special intensive therapy using neurostimulation. In the Netherlands, that technique was considered experimental and only to be used in certain circumstances. Pursuant to guidelines operating in the Netherlands, Peerbooms would not have qualified for such experimental treatment due to his age. Thus the Sickness Fund refused to pay the costs of treatment.

In both cases, the ECJ had to rule on the pre-authorization rule, and whether such a rule in the circumstances of those particular cases constitutes a barrier to the freedom to provide services. In contrast to the Decker and Kohll ruling (dealing with non-hospital care within a reimbursement system), the Smits/Peerbooms case concerned access to hospital care services for which the sickness fund had not contracted, and which, within the Netherlands, is provided on a “benefit-in-kind” basis. The Court agreed that sickness funds should not be exposed to the cost of hospital services for which they had not contracted. However, the pre-authorization condition as applied by the authorities in the Netherlands was criticized for its potentially discriminatory effect. In the Netherlands, the general legal rule under which the costs of medical treatment are covered is where the treatment is found to be “normal in the professional circles concerned.” This expression, however, is open to a number of interpretations, depending in particular on whether what is “normal” is considered as such in Dutch medical circles (this narrow interpretation being favored by the national court in the Netherlands). In contrast, the ECJ decided that to allow only treatment habitually carried out on national territory and scientific views prevailing in national medical circles to determine what is or is not “normal,” will not offer sufficient guarantees to patients that the guidelines in place are objective, non-discriminatory, known in advance and not used arbitrarily. Moreover, such a focus on national conceptions of “normal” will make it likely that Dutch providers will always be preferred in practice. The ECJ found that where treatment is sufficiently tried and tested by international medical science, refusal of the prior authorization cannot be justified. Further, to satisfy the “normal” criterion, a Member State “must take into consideration all the relevant available information, including, in particular, existing scientific literature and studies, the authorized opinions of specialists and the fact that the proposed

What has become clear from the Leichtle case is that national rules that restrict contracting to health institutions in the Member State exclusively are forbidden.

From this case, it became clear that Member States must apply the pre-authorization procedure consistently and that patients cannot be denied health care abroad arbitrarily (i.e., there have to non-discriminatory, transparent procedures, and appeal mechanisms). For patients entitled to benefit-in-kind services, such as provided in the Dutch system, the ruling means that it should be just as easy to receive medical treatment from a foreign non-contracted provider as it is to obtain from a non-contracted provider in the country of insurance. As such, the ECJ’s interpretation of communal pre-authorization conditions creates new opportunities for extended access to health care abroad. Subsequent to the Smits/Peerbooms case, the ECJ had to rule on two more or less identical situations in the Dutch mixed case of Müller-Fauré/Van Riet. Here, the Court consolidated and clarified its previous reasoning on prior authorization, at least concerning inpatient hospital care. However, the Court also confirmed that there are several reasons that may justify requiring prior authorization where social health insurance funds cover benefits provided in another Member State. These reasons include: the protection of public health in as much as the system of agreements is intended to ensure that there is a high-quality, balanced medical and hospital service open to all; to guarantee the financial balance of the social security system; and, finally, to enable managing authorities to control expenditures for and the planning of health services. The Court noted that concerns regarding undermining
the financial balance of the social security system particularly valid vis-à-vis hospital care. In the case of hospital services, according to the Court, it is well known that to ensure sufficient access to a wide range of hospital services and in order to contain costs, that careful planning is required regarding the number of hospitals, their geographical distribution, the mode of their organization and the equipment with which they are provided. Nonetheless, the conditions attached to the grant of authorization must be justified and satisfy the requirement of proportionality, and such a prior authorization scheme must likewise be based on a procedural system which is easily accessible and capable of being challenged in judicial or quasi-judicial proceedings.

However, with respect to outpatient (non-hospital) health care, as is the case in Müller-Fauré (i.e. dental care), the Court was not convinced that abolishing prior authorization will have a system-undermining effect. According to the Court, “there was no evidence that indicated that the removal of the prior authorization requirement for that type of care would give rise to patients travelling to other countries in such large numbers, despite linguistic barriers, geographic distance, the cost of staying abroad and lack of information about the kind of care provided there, and that the financial balance of the social security system would be seriously upset.” Therefore, in case of non-hospital services, there was no justification for requiring prior authorization when one applies the free movement principle.

Although the Müller-Fauré/Van Riet case was considered the latest in a line of cross-border rulings, new Court rulings on hospital-related expenditures in another Member State (the Leichtle case), as well as the reimbursement of hospital services in non-EU countries, have further extended the notion of patient mobility (the Keller case). Although the cases question the German (Leichtle), and Spanish (Keller) legal norms on cross-border care, the outcome is relevance to all Member States, including the Netherlands.

In the Leichtle case, the question raised did not concern so much the approval and reimbursement of the expenditures of the health care facility abroad (i.e., health spa), but the rules concerning the reimbursement of other expenditures related to the treatment abroad (travel, lodging, etc.). Since the conditions for reimbursement of this kind of costs were set for out-of-country treatment only (i.e., increased prospects of success, and the report written by a medical officer), Germany could deter the insured from approaching health care providers abroad, ergo, hindering free movement. Expenditures related to board and lodging can be considered as an integral part of the health care itself. After all, just as hospital treatment may involve a stay in hospital, a health care services administered for therapeutic purposes may well, by its nature, involve admission at a spa. Although travel costs and visitor’s tax are not medical in character, they are, according to the Court, “inextricably linked to the cure itself since the patient is required to travel and stay at the spa in Italy.” Consequently, the conditions for these expenditures have to be tested according to the previously accepted reason for justification. Additionally, this means that the measure taken should be necessary and that its objective cannot be reached by an alternative, less invasive measure under the same conditions (proportionality test).

The German authorities claimed that the absence of the disputed conditions would seriously harm the financial equilibrium of the German social security system if it is not accompanied by an analysis of the appropriateness and proportionality of the restrictive measure. Since they could not support that claim with well-founded arguments, the Court did not accept the general-interest reason as justification for restricting the free movement of patients. Therefore, Mr. Leichtle was compensated for the additional expenditures of the health spa abroad. A contrario, in case a Member State can show there is no less intrusive alternative equally effective as the suggested measure, or make a reasonable case for the financial instability of the system in case many people would access health spas abroad, the outcome of this case may have been different.

What has become clear from the Leichtle case is that national rules that restrict contracting to health institutions in the Member State exclusively are forbidden. However, the condition that a treatment should be provided by an institution listed in a so-called “Registration of Health Spas” does not necessarily hinder access to spa services in another Member State, since the rationale of such a measure is to ensure that sickness funds can check the “seriousness” of services provided by health spas, in and outside the country. Nonetheless, the registration requirement may still have a potential hindering effect, which depends on the objectivity of the conditions for registration. Finally, the definition of the health plan entitlements, as well as the amount of reimbursement granted, remains the prerogative of the member states themselves. This is caused by the absence of harmonizing competences at Community level in the field of social security.

Most recently, in the Keller case, the Court was confronted with the question whether EU citizens, may ask for (the reimbursement of) medical treatment outside the European Union based on European social security rules. More specific, the Court was being asked whether the granted authorization, provided by the Spanish health authorities (Insalud) to undergo a med-
tical treatment in a Member State (Germany), should be interpreted as an approval for a surgical operation that, in view of its extreme delicate nature and the special expertise required, could only be performed in a non-Member State private clinic (Switzerland). In other words, whether the Insalud is bound by the diagnosis and choice of treatment of the doctors authorized by the German medical bodies, and thus whether consequently there is an obligation to reimburse the costs of Keller’s hospital treatment the Swiss clinic.

The Court refers to the objective of the applicable social security provision, and its function, i.e., helping to facilitate the free movement of persons covered by social insurance, and to the same extent, the provision of cross-border medical services between Member States based on Article 22 of regulation 1408/71. The achievement of the objective pursued is based on a sharing of responsibilities between the competent institution (Insalud) and the institution of the Member State of stay (Germany). Whereas the Insalud is responsible for granting authorization for receiving health care abroad, it is for the German physicians to provide those services in accordance with the applicable German law. It follows from that rule of shared responsibilities, in correlation with the Union measures relating to the mutual recognition of diplomas of the formal qualifications of practitioners of medicine, that the Insalud is bound by the hands of the German physicians, and that it is obliged to accept and recognise the findings and choices of treatment made by those doctors. In this particular case, Insalud gave permission to undergo a medical surgery in Germany. Accordingly, when it has agreed to receive a medical treatment outside Spain, it is bound by the findings relating to the need for urgent vital treatment made by the authorized doctors. Similarly, Insalud is bound by the choice of treatment made by these doctors given the circumstances that the urgent treatment could only be provided outside the EU.

Under these circumstances, Ms. Keller cannot be required to return to the competent authority, the Insalud to undergo a medical examination there, when doctors in Germany considered that her state of health required urgent treatment. Moreover, asking prior approval of the Insalud would disregard the rule of shared responsibilities underlying European social security law and the principle of mutual recognition of doctors’ professional skills.

With respect to the conditions for reimbursement of the costs linked to medical treatment in a non-Member State, the Court refers to its previous rulings in which it concluded that, in principle, the insured enjoy the benefits-in-kind provided by the health services of the Member State of stay, on behalf of the competent authority. In this particular case, it means that the costs of treatment borne by (the heirs of) Keller have to be reimbursed by the Insalud, in accordance with German reimbursement rules. By doing so, the Spanish insured are treated under the same conditions as insured citizens within Germany. Furthermore, it should be clear that the treatment received is among the benefits provided by the legislation of the competent member state, i.e., Spain.

The Keller case can be added to a series of rulings from the European Court of Justice on health care abroad. The difference with other rulings is that Keller asked for (reimbursement of) medical treatment outside the European Union, whereas previous rulings concerned in and out patient health care in another Member State. What makes this case interesting is the interpretation given by the Court on European social security law, more specific the meaning of Regulation 1408/71 on the coordination of social security. Although the territorial scope of Union law is restricted to its Member States that does not exclude its applicability when it concerns services provided outside the European Union. Essentially, this is what the Court of Justice has made clear. The Spanish government disputed such an extensive interpretation of Community law by claiming that since the medical treatment provided outside the European Union, the outcome of the proceeding does not therefore depend on the interpretation of Union law, but is a matter exclusively for national law. The Court parried that argument by stating that the decisive criterion for the applicability of the regulation is that the insured is affiliated to a social security scheme of a Member State. Since Ms. Keller was affiliated under the Spanish Insalud, the applicability of Union law was therefore beyond doubt.

Furthermore, the Court based its decision on the rule of shared responsibilities between the home Member State and the Member State of stay. Secondly, due to the system of mutual recognition of diplomas, one may assume that German doctors are qualified and authorized to make a diagnosis and choice of treatment. Given that competence, the fact that the state to which the patient has been transferred is not a member of the European Union, is of no importance.

4.2 Consequences for the Dutch Health Insurance Scheme
One may question whether the Court’s case law, and in particular the recent ruling in Keller, has not opened a Pandora’s box. Will the result be that patients, now empowered to search for the best health care services abroad, begin the eventual demise and dismantlement of the social health insurance system? The answer is, in short, no; but the implications for the Dutch – and other – health care schemes are considerable.
First, it is now clear that medical activities fall within the scope of the EC treaty provision on free movement of services, even hospital services. Consequently, national social security rules cannot be used to exclude application of the free movement provision. This applies to insurance schemes, such as in the Netherlands, providing benefits-in-kind services, but also to hospital services provided by a National Health Service such as exists in the UK. The UK government unsuccessfully attempted to exempt NHS services from the ambit of article 50 of the EC since it provides services directly rather than reimbursing the cost of services received.

Hospitalization should be interpreted in the patient’s favor (as opposed to the national authorities’), meaning that the place of actual treatment should be decisive in interpreting whether hospitalization is required.

The ECJ noted “that a medical service does not cease to be a provision of services because it is paid for by a national health services or by a system providing benefits-in-kind.”

Second, the Court has also, on occasion, made it clear that EU law does not undermine the power of the Member States to organize their respective social security systems. In the absence of harmonization at EU level, each Member State may pass legislation pursuant to which citizens have first a right or duty to be insured with a social security scheme, and, second, the conditions for entitlement to benefits. Moreover, it is not incompatible with EU law for a Member State to establish, with a view toward achieving its aim of limiting costs, a negative list excluding certain products from reimbursement. It follows that EU law cannot, in principle, have the effect of requiring a Member State to extend the list of medical services paid for by its social insurance system. Nonetheless, in exercising its powers, the Member State must not disregard EU law. Essentially, this means that the list of insured medical treatments must be drawn up in accordance with objective criteria, which are known in advance, and without reference to the origin of the service (non-discrimination). In the Netherlands, the health insurance system is not based on a pre-established list of types of treatment for which payment will be guaranteed; rather the legislature has enacted a general rule providing that all costs of medical treatment will be assumed, provided that the treatment is “normal in the professional circles concerned.” It is thus largely up to the discretion of the social health insurance funds to decide which types of treatment satisfy that condition; however, in applying that criterion, these funds must now interpret the criterion on the basis of what is sufficiently tried and tested by international medical science. This could mean that where a certain treatment has sufficiently been tried and tested by international science, authorization by the sickness fund could not be refused on the grounds that it is not presently provided in the Netherlands. The only justifiable reason to refuse approval is where, given the need to maintain an adequate supply of hospital care and to ensure the financial stability of the sickness insurance system, the “same or equally effective treatment can be obtained without undue delay” at a contracted provider. It should be noted that in determining whether “the same or equally effective treatment can be obtained without undue delay,” the mere fact that a person is on a waiting list does not necessarily mean that the treatment is unavailable. “Undue delay” should be determined as the period within which medical treatment is necessary with respect to the patient’s medical condition and history.

Moreover, those Member States, like the Netherlands, that use the “normal criterion” in determining entitlements for coverage will not be required to allow their citizens to obtain treatment in any hospital in any EU country unless that treatment is both accepted in international medical circles and is not sufficiently available in the home country (medical necessity criterion).

Third, patients are now entitled to search for non-hospital services abroad that are not available in their home country or, if available, are not available in a timely fashion (Müller-Fauré/Van Riet). One should keep in mind, however, that without prior authorization, social health insurance funds are still fully entitled to reimburse only costs up to the maximum amount that is applicable in the country of residence. For that reason, the Court did not consider the removal of the administrative prior authorization condition to be a serious threat to the financial balance of the social security system, since it had to bear the cost of treatment when received in the patient’s home country anyway. Patients seeking non-hospital care in other EU nations without prior authorization must bear any additional costs above and beyond the relevant tariff that provided for in the relevant tariff in their own country.

Fourth, as mentioned above, the larger public interest in maintaining a sustainable social insurance system may be accepted as justifying barriers to freedom to provide medical services in the context of hospital infrastructure. Member States need to determine whether their respective national rules can be legitimately justified in the light of such overriding reasons.
In accordance with settled case law, it is necessary to ensure that they do not exceed what is objectively necessary for the given purpose and that the same result cannot be achieved by less restrictive rules. As determined in the Müller-Fauré/Van Riet case, these requirements apply regardless of the type or nature of the health care system (e.g., whether a social insurance system as in the Netherlands or a national health service as in the U.K.).

Fifth, the various ECJ rulings leave some unsettled questions particularly given that the difference between hospital (inpatient) and non-hospital (outpatient) care is not always that clear. For example, some surgical services may be provided in a hospital or in an outpatient clinic. Moreover, certain types of care are only partly hospital-based, e.g., hospital treatment combined with admission to an outpatient clinic. Disagreement about the nature of the provided care concerned may give rise to legal uncertainty among the insured and cause an increase in litigation. Patients need to be particularly aware of whether or not the provided services will be classified as a hospital service. In such a case if a patient has not obtained prior authorization before receiving treatment in another Member State, then he or she may be denied reimbursement.

In a recent ruling, the ECJ confirmed the nexus between the need for prior authorization and hospitalization finding that where the “multidisciplinary treatment of pain which the claimant envisages […] involves her hospitalization, then this necessitated the patient obtaining prior authorization.” In contrast, the District Court in Maastricht concluded that a special type of physiotherapy, which requires hospital admission, should not be considered to be hospital care, since it is generally qualified as non-hospital care.

These contrasting cases demonstrate that the difficult question of what constitutes hospital care or non-hospital care is decisive for determining whether prior authorization is required and thus whether patients will be reimbursed for the costs of treatment received in other Member States. However, in some cases there may be no communis opinio among medical professionals about whether or not hospitalization is necessary. In that case, patients searching for alternative treatment options in other Member States without prior authorization take a considerable financial risk. Arguably at least, hospitalization should be interpreted in the patient’s favor (as opposed to the national authorities’), meaning that the place of actual treatment should be decisive in interpreting whether hospitalization is required. Any other interpretation may create a perverse incentive for social insurance funds to organise outpatient care in a hospital sphere for merely opportunistic reasons.

Sixth and finally, based on the Court’s extensive interpretation of the patient mobility provision in the Keller ruling, in exceptional cases, the sickness funds can be forced to reimburse the costs of treatment in a country outside of the EU. However, as a rule, the treatment provided, should be recognised by the Member State of stays’ national law.

In sum, there is no doubt that the Court’s jurisprudence on patient mobility has significantly contributed to the development of a corpus of rights to access health care, although one may question whether the Court has gone too far at the battlefield of cross-border care. In contrast to Field Marshal Montgomery’s experience with World War II’s most tragic blunder, Operation Market Garden at the bridge of Arnhem, the Court has succeeded to force a break in the Dutch and other Member States’ line of defence on social security. Starting with the Decker and Kohll ruling, further case law has extended the patient’s options to receive health care abroad and simplified the necessary conditions, and simultaneously, taking into account the public interest. In this respect, the Keller case was definitely not the last battle in this area. By strengthening the rights of patients, there is no bridge too far.

5. Latest Developments on Social Health Insurance Reforms

Recently, the Parliament has approved a new Health Insurance Act, introducing cost-sharing measures that may affect high-risk persons disproportionally, a change towards a more restricted notion of solidarity is emerging.

There are, however, also reasons for concern. Generally interpreted as a social security right in the Netherlands, access to health care is based on the concept of solidarity and equality. With the new Health Insurance Act, introducing cost-sharing measures that may affect high-risk persons disproportionally, there are also reasons for concern.
Although the “Zorgverzekeringswet” has been inspired by both international and European law, critics have questioned the law’s conformity with treaty law.\textsuperscript{72} So far, that debate was primarily focussing on the new role of health insurance entities when introducing a dose of competition in health insurance in relation to EU law. However, little attention has been paid to the international legal consequences of the introduction of cost-sharing measures, such as the no-claim refund and compulsory co-payments. The no-claim refund should encourage a less health care-consuming attitude and, \textit{ergo}, an incentive to cost containment in healthcare. In return, the insured will receive a yearly premium reduction when the costs of insured care remain below a certain limit (€255). Although it may achieve the goal of cost reduction there may be serious consequences for certain categories of patients (e.g., chronically ill) that will not benefit from that measure since they will certainly exceed that yearly limit. According to the governmental advisory body, the Council of State (“Raad van State”), the no-claim refund should be considered as a cost-sharing measure and thus may fall within the material scope of international social security treaties, such as ILO Convention 102 and the European Social Security Code.\textsuperscript{73} Although the no-claim refund excludes certain services from the threshold, such as GP-care, maternity and delivery care, it is not clear whether these exemptions will be sufficient from the treaty point of view.\textsuperscript{74}

As a rule, cost-sharing should be so designed as to avoid hardship, but considering the no-claim refund, in combination with the introduced care-related co-payments and the nominal premium, there is a serious risk that the total costs will exceed the specified percentages as set in the European Code (protocol), and therefore violate the hardship-clause.\textsuperscript{75} In a way, history may repeat itself given the Central Appeals Tribunal’s ruling prohibiting co-payments for maternity care.\textsuperscript{76}

The hardship-clause can be considered the Code’s equivalent of the “standstill” provision, as it is known in international treaty law. In the International Covenant on Economic, Social and Cultural Rights (ICESCR), the standstill clause has been defined as “[a]ny deliberately retrogressive measure would require the most careful consideration and would need to be fully justified by reference to the totality of the rights provided by the Covenant and in the context of the full use of the maximum available resources.”\textsuperscript{77} Although the standstill-clause is not absolute, “there is a strong presumption of impermissibility of any retrogressive measure,” explained by the State party.\textsuperscript{78} Also in that respect, a retrogressive measure such as the no-claim refund arrangement, in combination with co-payments and a flat-rate premium, is highly questionable, and can only be justified by means of the “healthcare allowances” to be introduced simultaneously with the “Zorgverzekeringswet.”\textsuperscript{79} But in case such a grant appears insufficient, and the financial burden hinders the effectiveness of medical protection to (high-risk) persons, the impermissibility of the cost-sharing measure revives, causing a violation of the Treaty’s core obligation.\textsuperscript{80} Whether or not the judiciary would accept such a hardship or standstill argument remains to be seen. First of all, it requires hard evidence of the retrogressive effects of the new law.

6. Final Remarks

It is clear that international law, including European Union law, has had a significant influence on guaranteeing access to healthcare in the Netherlands. International legal norms conceptualise the content of the health care right, whereas the judiciary confirmed the enforceability of such a right in terms of benefit entitlements. Moreover, the European Court of Justice, with its rulings on patient mobility has extended the meaning of such a right, in the Netherlands and abroad.

There are, however, also reasons for concern. Generally interpreted as a social security right in the Netherlands, access to health care is based on the concept of solidarity and equality. With the new Health Insurance Act, introducing cost-sharing measures that may affect high-risk persons disproportionally, a change towards a more restricted notion of solidarity is emerging. Although the government claims that the new Health Insurance Law does not alter the nature of the social health insurance scheme, the competitive and for-profit nature of the insurance and providers market, or the increased elements of capitalisation (e.g., the no-claim refund, medical care saving accounts considerations, etc.), the evidence is otherwise. It reflects a gradual shift from social towards direct insurance that seems difficult to match with the solidarity notion, at least in terms of European and international law. The development of liberalizing the health (insurance) market challenges the country’s autonomy to define the entitlements, as well as setting conditions restricting access to healthcare. A simple referral to the financial sustainability clause will not be sufficient to convince the judiciary of the importance of access’ restrictive measures. As such, international treaty law challenges the future sustainability of the system and thus, ironically, may result in retrenchment in the entitlements as seen in the Netherlands.

References


3. As an exception to that rule, the courts may review the constitutionality in terms of international and European law.

4. According to Article 22(1) of the Dutch Basic Law the authorities shall take steps to promote public health. This provision has been generally interpreted as a “mere” obligation of the government to be concerned with setting up health facilities and facilitating access to necessary healthcare. A. P. den Exter and H. G. M. Hermans, o.c. (1999): 271.

5. Ziekenfondsact of October 15, 1964, Bulletin of Acts, Orders and Decrees 1992, 391, as most recently amended by the Act of December 23, 2004 (Bulletin of Acts, Orders and Decrees 2004, 725). There is a universal scheme covering the whole population, the so-called “AWBZ”-scheme covering exceptional medical expenses. Beyond this citizens who don’t are over the income threshold for the Ziekenfonds may be covered by a private insurance standard policy for special categories of insured (based on the Health Insurance Access Act [WTZ 1995]) and/or voluntary private insurance schemes.

6. The conditions differ by type of care and include prescription by a general practitioner, quantitative maxima for prescribed pharmaceuticals and prior approval by the insurance fund.

7. As an exception to that rule, the courts may review the constitutionality in terms of international and European law.

8. E.g., International Labour Organisation (ILO) Conventions concerning minimum standards of social security and maternity protection (102 and 103), the International Covenant on Economic, Social, and Cultural Rights (ICESCR), the Council of Europe’s Social Charter (ESC) and the European Code on Social Security.


11. The Central Appeals Tribunal concluded that national regulation that imposed cost sharing in case of maternal assistance was a violation of Article 10 of the ILO Convention 102 setting minimum standards on social security (CRvB May 29, 1996, RSV 1997/9). According to the Tribunal the imperative nature of the treaty obligation (obligation to ensure), and the concrete wording of the treaty provision prohibit the introduction of cost sharing measures. As a consequence of this ruling, the Dutch government annulled the Decreé introducing own contributions for hospital admissions.


13. E.g., In General comment no. 3 on the nature of States Parties obligations (Art. 2(1) of the ICESCR). UN Doc. E/1991/23, para. 5, the Committee claims that: “any suggestion that the provisions indicated are inherently non self-executing would seem to be difficult to sustain”. General comment no. 14 on health. The right to the highest attainable standard of health. UN Doc. E/C.12/2000/4, para. 30; UN Doc. E/C.12/1/Add.25.


15. Supra note 14, Article 8 of the Convention which, in so far as relevant, provides as follows: [...] everyone has the right to respect for his private [...] life. There shall be no interference by a public authority with the exercise of this right except such as is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety or the economic well-being of the country [...].


17. E.g., Stubbings and Others v. the United Kingdom, judgment on October 22, 1996 (Reports 1996-IV): 1501, 62.

18. See, Nitecki v. Poland, judgment on March 21, 2002, App. No. 56563/01 §1, where the Court considered that: “it cannot be excluded that the acts of omissions of the authorities in the field of healthcare policy may in certain circumstances engage their responsibility under Article 2.”

19. The new EU Constitutional Treaty will replace the previous Treaty of the European Union incorporating previous amendments, OJ 2004/C 310/01. However, since the French and Dutch rejection of the Treaty by referendum (June 2005), it is unlikely that this new Treaty will come into force.


22. Inter alia, articles 94 and 308 EC. Both provisions are aimed at the realisation of the common market. Further restriction of intra-communautaire trade however, can be found in other relevant treaty provisions aimed at the protection of health and life of humans, for instance article 30 EC.

23. And therefore, by reason of the scale or effects of the proposed action, be better achieved by the Community, Article 5 EC.


26. Article 152(4)(a)(b) EC.


28. E.g., the mutual recognition directives, which facilitate a general system for the recognition of diplomas of certain health professionals (Directives 89/48/EEC and 92/51/EEC).


30. Art 234 EC.

31. Deckers v. Caisse de Maladie des Employés Privés C-120/5 (1998) ECR I-1831 and Kohll v. Union des Caisses de Maladie C-158/96 (1998) ECR I-1931. The European Court of Justice overruled the pertinent Luxembourg regulations, which made reimbursement by the social security system of medical services provided in another Member State – respectively orthodontic treatment and the supply of spectacles – conditional on prior authorization. The Deckers and Kohll ruling has initiated a number of cases questioning the Dutch health insurance system, e.g., Geraerts-Smith/Peerboms C-157/99 (2001) ECR I-5473; Müller-
André den Exter

This means, for instance, to include evidence from clinical randomized controlled trails (stage III studies), see, Health Care Insurance Board (CVZ), February 23, 2005 (RZA 2005): 37; District Court Utrecht, December 8, 2004 (RZA 2005): 38.

35. Smits/Peerbooms, supra note 31, para. 103.

36. Here, the Central Appeals Tribunal refers to standards, so-called "Treek-normen," defining the maximum acceptable waiting time for a specific medical intervention. The underlying idea is that hospitals, for planning and efficiency reasons, need a certain waiting time. CRvB June 18, 2004, para. 92. Also, CRvB June 18, 2004, USZ 2004/277, 11 month waiting time for a hip replacement is not considered as "timely."


38. Kohll, supra note 31, para. 41; Müller-Fauré/Van Riet, supra note 31, para. 72.


40. As a rule, the Health Care Insurance Board (CVZ) decided that hospital care requires at least one day admission in an hospital institution. Letter no 03/35, June 25, 2003.

41. Also suggested by Advocate General Colomer in his opinion on Smits/Peerbooms, supra note 31, para. 61. The risk of denial of reimbursement is also present in case a non-hospital treatment still requires hospital admission, for instance due to medical complications. In such a case, since prior authorization is absent, the costs of hospital admission and treatment will not be reimbursed.


43. Maastricht District Court, September 26, 2003, AL3183, the Netherlands.


45. A similar interpretation – mutatis mutandis – was used by the Court in the Vanbraekel case, supra note 47, when ruling that the level of reimbursement for foreign hospital services should have been the same had the same services been received in the home country.

46. Operation Market Garden was aimed at crossing the bridge of Arnhem, near the German borders, in order to march on to Berlin. Unfortunately, British paratroops were dropped behind enemy lines, whereas allied forces could not release them since the German army occupied the bridge over the river Rhine. This historic battle inspired Hollywood filmmakers to shoot a film titled "A Bridge too Far."

47. The “Zorgverzekeringswet” (2001) was approved by Parliament on June 14, 2005 and will come into force on January 1, 2006; (Official Journal of the European Union: 35, 2005).


49. E.g., Kohll ibid. para. 18. In the Dutch situation, the nature and scope of entitlements are being set by the main social insurance scheme, the Sickness Fund Act ("Ziekenfondswet").

50. Duphar and Others ibid. para. 16.

51. Duphar and Others ibid. para. 16.

52. This means, for instance, that they should respect basic community principles, including the non-discrimination principle. Ferlini v. Centre hospitalier de Luxembourg C-411/98 (2000), ECR I-8081. In this case, the Court ruled that the application, on a unilateral basis, by a group of healthcare providers to EU officials of scales of fees for medical and hospital maternity care that are higher than those applicable to residents affiliated to the national social security scheme constitutes discrimination on the ground of nationality prohibited under Article 12(1) EC, in the absence of objective justification.

53. Verstrekkingenbesluit ziekenfondsverzekering (Sickness Funds Decree on Entitlements article 3(1); Huisartsenverzorg ten verlenen door een huisarts omvat genees- en heelkundige zorg naar de omvang bepaald door hetgeen in de kring der beroepsgenoten gebrui-kelijk is,…(GP care includes medical care as determined by what is normal in the professional circles concerned); and mutatis mutandis article 12(1)(a), and 13(1).

54. This means, for instance, to include evidence from clinical randomized controlled trails (stage III studies), see, Health Care Insurance Board (CVZ), February 23, 2005 (RZA 2005): 37; District Court Utrecht, December 8, 2004 (RZA 2005): 38.

55. Irrespective the suggested "healthcare allowance," since only a limited number of insured will profit this allowance, an incomerelated contribution to the cost of the nominal insurance premium that will be introduced with the "Zorgverzekeringswet" simultaneously.

LEGISLATING AND LITIGATING HEALTH CARE RIGHTS AROUND THE WORLD • WINTER 2005 709
76. See also, M. W. Wettens-Bronsgeest. “Eigen bijdragen aan een zijden draadje?” (“Own payments by a threat, in Dutch”), Rechtspraak Zorgverzekering, RZA (2004): 979-980. The European Code corresponds to ILO Convention 102, but foresees in (maximized) co-payments but avoiding hardship.


79. Under the new regime, the insured will pay a nominal insurance premium to their health insurer. To keep the health insurance system financially affordable for all, a “healthcare allowance” will be introduced. But that allowance will not compensate premium cost entirely.

80. Concluding Observations Iraq UN Doc. E/1998/22, at 253, in which the UN Committee concluded that economic sanctions imposed by the UN on Iraq do not justify retrogressive measures towards Iraq’s treaty obligation to guarantee food and pharmaceuticals. Although the standstill clause is not absolute, see: M. M. Sepúlveda Carmona, The Obligations of the State under the International Covenant on Economic, Social and Cultural Rights, (Utrecht: Utrecht University [diss.], 2002): 348-349.
Historically, judicial enforcement of constitutional rights to health care has played a fairly limited role in enabling access to health care, a trend particularly prevalent in North America, and reflected in many other regions. This trend is due in part to judicial resistance to recognizing socioeconomic rights like health as appropriately legal, or as appropriately enforceable in light of the doctrine of separation of powers. This resistance is evident in judicial deference to social and economic policy, a reluctance to view socioeconomic claims as invoking “fundamental values” that courts consider themselves authorized to protect, and a real reluctance to recognize and enforce “positive” obligations pertaining to social welfare. As a result, health has often fallen largely into the political rather than legal sphere, and domestic courts have been relatively reluctant to review health policies from a human rights perspective, given the belief that doing so would exceed the appropriate democratic function of the judiciary.

International law both refutes and supports these domestic perceptions. International law has long recognized health as a human right, and health rights entrenched in a wide array of international and regional human rights instruments undermine arguments that health is not an appropriately “legal” human right. Similarly, arguments that treaty rights (such as the right to health in article 12 of the International Covenant on Economic, Social and Cultural Rights [ICESCR]) are indeterminate and vague have been put to rest by authoritative international interpretations of the entitlements and duties that such rights impose. Yet while international human rights law holds as a foundational article of faith that social rights have equal and indivisible importance to civil rights, an international enforcement mechanism for ICESCR rights has yet to be established, a lacuna which reinforces perceptions that these rights are not appropriately justiciable. This disjuncture between the entrenchment of health rights and their enforcement is reflected at the domestic level. While a recent study indicates that over two thirds of all constitutions have provisions regarding health and health care, the dearth of global health-rights jurisprudence suggests that these rights have not enjoyed widespread judicial enforcement.

This picture is slowly changing, and there is a small but growing jurisprudence on health-related rights, led in particular by the South African Constitutional Court’s enforcement of the domestic constitutional rights provisions. The South African Constitutional Court has been at the forefront of asserting the human rights to health and to health care, and has consistently held that health and health care are fundamental rights that governments are required to fulfill. This jurisprudence has been influential in shaping the approach to health rights in other countries, and has helped to establish health as a human right that is justiciable in domestic courts.

Lisa Forman, LL.B, is a doctoral candidate at the University of Toronto, Faculty of Law, researching the role of human rights in relation to accessing AIDS medicines. She received a BA and LLB from the University of the Witwatersrand, and an MA (Human Rights Studies) from Columbia University.
right to access health care services. The nation's unfolding judicial and legislative response to this right is thus uniquely illustrative of its potential impact on claims for health care through the courts, on health policy-making, and on legal adjudication itself. The South African Constitutional Court's adjudication of health rights goes some way towards subverting the notion that enforcing social rights exceeds judicial capacity and violates the appropriate democratic separation of powers between the three branches of government. However, the South African experience also underscores the importance of the individual and institutional choices made by judges in enforcing health rights. Although the constitutional entrenchment of health rights has ensured some degree of justiciability, the Court retains discretion regarding its standard of judicial review of socioeconomic policies, the extent of its deference to policy deficiencies and resource allocations, the scope and content of the entitlement and duties these rights confer, and the nature of the relief it would order for litigants.

The transformative potential of these rights in relation to health and health care largely depends on the nature and quality of the choices the Court makes in relation to these questions. However the traditional objections to health rights still have some purchase in the Constitutional Court's approach to these rights, reflected in former Chief Justice Chaskalson's view that these rights sit at "the border of the separation of powers between the judiciary and the executive."8 The Court's attempts to balance claims for remedial action with what it considers constitutionally appropriate deference therefore frames the exercise of its interpretive discretion of these rights, with mixed results for advancing their adjudication.

This article explores the impact of South Africa's constitutional health rights on judicial decision-making in the context of relevant jurisprudence, national health needs and health policy. Despite a government rhetorically committed to transformation (the meta-slogan of post-apartheid governance) and to increasing access to health care, progress has been fairly limited and in cases such as HIV/AIDS treatment policies, actually regressive. Despite a government rhetorically committed to transformation (the meta-slogan of post-apartheid governance) and to increasing access to health care, progress has been fairly limited and in cases such as HIV/AIDS treatment policies, actually regressive.

1. Constitutional Health Rights in Context

South Africa's racist history under apartheid pervades the choices made in its new democratic Constitution, expressly adopted so as to “[h]eal the divisions of the past,” and to establish a “democratic and open society in which government is based on the will of the people and every citizen is equally protected by law.”9 Accordingly, the Constitution's foundational commitment is to creating an open and democratic state based on equality, dignity and freedom,10 a model of governance diametrically opposed to the apartheid state. Yet the Constitution's aspirations are also to “improve the quality of life of all citizens and free the potential of each person.”11 These intentions are manifest in the startlingly progressive rights protected in the 1996 South African Constitution,12 and in particular the entrenchment of a range of justiciable social and economic rights.13 The inclusion of these rights was a direct response to the gross poverty and extreme inequality that apartheid's material redistribution had produced for black South Africans.14 Although classified as a middle-income developing country, South Africa has a gross domestic product comparable to far wealthier countries.15 However, South Africa's wealth distribution is one of the most unequal globally,16 with poverty levels of approximately forty percent, high unemployment and many households experiencing limited access to education, health care, electricity and clean water.17 Despite the growth since 1994 of a black middle class and affluent sector, the vast majority of poverty remains concentrated amongst black South Africans.18 Addressing poverty, inequality and inaccessible basic needs therefore posed a tremendous challenge to the new democratic government, made considerably more
difficult by the democratic government’s inheritance of apartheid’s intentionally deficient public infrastructures.

1.1. Health Rights in the Constitution
In addition to rights to adequate housing, food, water and social security,20 the Constitution entrenches a range of health-related rights. These include children’s rights to basic health care services, prisoner’s rights to adequate medical treatment, and the right to an environment that is not harmful to health or well-being.21 The primary health right is contained in section 27, which is formulated as follows:

1. Everyone has the right to have access to health care services, including reproductive health care;
2. The state must take reasonable legislative and other measures, within its available resources, to achieve the progressive realisation of each of these rights.
3. No one may be refused emergency medical treatment.

This focus on the provision of health care is a markedly narrower formulation than the international human right to the highest attainable standard of health.22 The constitutional right effectively excludes the government’s broader public health responsibilities, given the relatively limited contribution that health care makes to improving population health.23 Indeed social conditions such as poverty and inequality are far greater determinants of health, given the direct correlation between disease prevalence and socio-economic status.24 However, the Constitution makes no explicit provision for these broader public health responsibilities and this has important implications for the ambit of the right with regard to public health more generally.25

The scope of this narrower right to access health care services is not explicit from the wording of section 27, which provides little indication of what is encompassed within the entitlement to health care services, nor the extent to which resource limitations and progressive realization could permissibly limit the state’s duty to ensure access. Some guidance on the state’s obligations is provided by the constitutional mandate that the state must respect, protect, promote and fulfil the Constitution’s rights.26 Drawn from international human rights law this typology implies both positive and negative duties with respect to each right.27 However the precise content of these duties was unclear and the task of interpretation fell primarily to the Constitutional Court.

1.2. Democratic Peril and the Constitutional Inclusion of Health Rights
The scope of individual entitlements and government duties under section 27 goes to the heart of the democratic hazards that enforcing social rights is perceived to hold. Thus at the Constitutional Court’s hearings to certify the Constitution, objectors to the inclusion of socioeconomic rights argued that enforcing these rights would breach the doctrine of separation of powers since courts would be making social policy and dictating budgetary allocations to government.28 The Court rejected these arguments as a bar to enforcement, holding that all judicial review carries budgetary implications, that much judicial review makes social policy, and that enforcing social and economic rights was certainly not so different a task from ordinary judicial review that it would breach the separation of powers.29 In the Court’s opinion, these rights were at least partially justiciable and at a minimum could be negatively protected from improper invasion.30 Less clear, however, was how the Court would enforce the state’s positive duty to fulfil such rights, given the scarcity of judicial precedents and the Court’s own concerns about its institutional capacity to formulate complex socioeconomic policy.31

2. The Constitutional Court’s Interpretation of the Right to Access Health Care Services
Since 1994, the Court has handed down three major decisions directly related to socioeconomic rights – the Soobramoney case in 1998,33 the Grootboom case in 2000,34 and the Treatment Action Campaign case (TAC) in 2002.35 These cases illustrate the Court’s evolving interpretation of the Constitution’s social and economic rights and similarly, its choices regarding the judicial role in enforcement. The Court’s first socioeconomic rights decision in Soobramoney was not, however, an auspicious start, and appeared to suggest a limited justiciability for section 27, an unwillingness to enforce positive fulfilment, and an undue deference to state assertions of budgetary constraints.36 In Soobramoney, the Court rejected an ill man’s claim for renal dialysis from a public hospital which rationed access for patients with chronic renal failure whose health status was incapable of improvement. The Court reasoned that rationing expensive health care in conditions of scarce resources and poverty did not breach section 27. The Court did not review the state’s averment that there were no resources available to expand access to dialysis, at least partially because Soobramoney never contested this claim.37 Given overspending and constraints on renal clinics countrywide, the Court considered rationing to be a rational response which maximized access to those who would benefit most.38 The implication seemed to be that the right was aspirational rather than...
The Court settled on a different approach in the Grootboom decision, which dealt with a claim under the constitutional right to adequate housing by people evicted by the state from shacks set up on public land. In Grootboom, the Court established a higher standard of reasonableness for state compliance with constitutional duties, and extensively interpreted the state’s obligations to take measures to progressively realize access to services within available resources – an obligation common to both the housing right in section 26 and the health right in section 27. The reasonableness standard was subsequently applied to health care in the TAC decision, where civil society groups challenged government’s policy on preventing maternal transmission of HIV/AIDS. This case is discussed in greater detail in part four below.

2.1. The Choice of a Standard of Reasonableness
In Grootboom, the Court indicated that while reasonableness was determined on a case-by-case basis, given great poverty and the constitutional commitment to equality, dignity and freedom, the state’s primary obligation was to act reasonably to provide the basic necessities of life to those who lack them. While the Court in TAC acknowledged that government could not conceivably meet even all basic needs, its obligation nonetheless is to seek to meet these needs, acting reasonably to provide access on a progressive basis. The Court stressed that the interrelated nature of all rights had “immense human and practical significance in a society based on human equality, dignity and freedom,” and meant that the socioeconomic rights guarantees must provide people with a “right to reasonable action by the state in all circumstances with particular regard to human dignity.”

I. REASONABLE LEGISLATIVE AND OTHER MEASURES
The Court held that this phrase means that at a minimum, the state must devise a comprehensive and workable plan to meet its obligations, providing for all needs, including short, medium and long term needs as well as for crises. Thus legislation, policies, and programs that exclude “a significant segment of society” will be considered unreasonable. In seeking to ensure that the basic necessities of life were provided to all, the state had to focus in particular on the needs of the most vulnerable, especially the poor, and particularly those experiencing urgent and desperate needs. Thus exclusions from coverage would not be constitutionally reasonable “if the measures, though statistically successful, fail to respond to the needs of those most desperate.” This interpretation is critical because of the qualitative dimension it brings to limitations in access, indicating that measures could not ignore “the degree and extent of the denial of rights,” nor exclude the particularly urgent needs of people whose “ability to enjoy all rights therefore [was] most in peril.”

Reasonableness applies to all elements of governance; not only the content of legislation, programs and policies, but also their manner of implementation. For instance, programs should be balanced and flexible, with national government bearing the responsibility of ensuring sufficient laws and polices to fulfil their obligations. In Grootboom, however, the Court emphasized that its review of reasonableness gave the legislature and executive a broad discretion on “the precise contours and content of the measures to be adopted,” and a court’s inquiry would be limited to whether these were constitutionally reasonable, “not whether other more desirable or favourable measures could have been adopted or whether public money could have been better spent.”

II. PROGRESSIVE REALIZATION
While progressive realization recognizes that full attainment of everyone’s right to access health care services is not always immediately possible, the Constitution’s goal was that “the basic needs of all in society be effectively met” and this requires that “the state must take steps to achieve this goal.” Thus, “accessibility should be progressively facilitated” with legal, administrative, operational and financial hurdles examined and where possible, lowered over time. These steps should also be taken as expeditiously and effectively as possible and any deliberately retrogressive measures would need full justification in light of all the rights in the Constitution and available resources. In the TAC decision, the Court emphasized the need for urgent and timely action where people’s lives literally depend on timely access to health care services.

III. RESOURCES
The reasonableness of any policy cannot be separated from the level of resources available, and the state is not required to do more than its available resources permit. This means balancing the goal (of meeting all basic needs) with the means. Equally importantly, trying to meet all needs requires “adequate budgetary support by national government,” and requires the state to “plan, budget and monitor” their fulfilment. While the Court will leave the choice of measures to realize health and the allocation of budgets to government, the implications are that government had to ensure
that budgets were arranged to reasonably address all needs, with a particular emphasis on the impact of any failures to do so upon poor and vulnerable people or those experiencing urgent and desperate needs.

2.2. The Limited Scope of the Entitlement and Duty under Section 27

The standard of reasonableness has a number of consequences for the scope of the entitlement and duty under section 27, and for the Court’s willingness to provide substantive content to this right. The Court has stressed through the cases that in the context of poverty, limited resources and great needs, both the state’s obligations and the corresponding rights are limited.60 Thus, not all health care needs could be met, particularly those so expensive as to threaten the state’s ability to provide other needs.61 Similarly, rationing of health care is constitutionally permissible if based on principled criteria.62

While the limited nature of the state’s constitutional obligation under this right is made explicit in the formulation of the right, the reasonableness standard underpins that while claims for health services must be reasonable (and may therefore be limited), so too must the state’s refusal of claims be reasonable. How this applies to health rights in particular is partially alluded to in Soobramoney, where, Justice Sachs, in a concurring opinion argues that “health rights by their very nature are interdependent,” requiring that traditional rights analyses be adapted to adjudicate between competing rights bearers.63 He argued however that this balancing should not be seen as imposing limits but as defining circumstances in which the rights may most fairly and effectively enjoyed.64 In Soobramoney, the Court asserted that balancing competing interests would at times require government to “adopt a holistic approach to the larger needs of society rather than to focus on the specific needs of particular individuals within society.”65 Implicit in such an approach is the necessity of proportionality between the interests to be balanced, an approach reflected in the TAC decision discussed below.

The understanding that the entitlement under these rights is limited also underpins the Court’s rejection of argument made by amici curiae in both Grootboom and TAC that it should recognize an entitlement to a “minimum core” of socioeconomic rights.66 The minimum core is an idea developed first in international human rights law, to reflect a state’s prioritized obligations and provide a floor below which realization should not fall without persuasive justification.67 In Grootboom the Court reasoned that recognizing a minimum core required that the Court itself determine its content, a complex task requiring information that the Court lacked in general and in the case before it. In the Court’s opinion however, this task in any event was not necessary given that the “real question in terms of our Constitution is whether the measures taken by the state... are reasonable.”68 In TAC, the Court extended its argument that it lacked the competence to determine the core, to argue that determining the minimum core would breach the appropriate separation of powers since it would require the Court to decide how public revenues should most effectively be spent.69 While the Court acknowledged that determinations of reasonableness could have budgetary implications, these were “not in themselves directed at rearranging budgets.”70 This distinction, the Court held, was what enabled the judicial, legislative and executive functions “to achieve appropriate constitutional balance.”71 In the Court’s estimation, the Constitution contemplated a “restrained and focused” judicial role, “namely, to require the state to take measures to meet its constitutional obligations and to subject the reasonableness of those measures to evaluation.”72 The role of the judiciary was therefore to ensure reasonableness and to protect “the democratic processes...so as to ensure accountability, responsiveness and openness.”73 Enforcing the minimum core in the Court’s opinion would therefore impermissibly overreach its more appropriately “restrained” role in relation to socioeconomic rights. The Court’s approach to the right to health therefore can be characterized as one that limits both the scope of the right and the judicial role in enforcement. However as this paper will proceed to illustrate, while this limited approach certainly misses opportunities for advancing the realization of the right, it nonetheless has had a positive influence on national health policy.

3. How has the Government Fared in Delivering Health Care since 1994?

It is beyond the scope of this paper to fully assess the extent to which governmental measures taken since 1994 accord with the constitutional standard of “reasonableness.”74 I will, however, evaluate in broad terms the effectiveness of government action in the health sector by considering changes in selected indicators of health care access and population health.

To contextualize this evaluation it is important to consider the tremendous challenges government faces in redressing health inequities and disparities, both within and between public and private health care. At the advent of democracy, the apartheid-era public health care system was highly inequitable and deficient, characterized by fragmented service delivery, insufficient rural facilities, and highly limited access to health care services for women, children and farm workers.75 The government’s commitment to transforming these inequities and deficiencies has formed a
Viewed from the perspective of reasonableness, South Africa’s HIV/AIDS policies do not fare well.

substantial part of national health policy since 1994, dedicated to achieving health equity and meeting the needs of vulnerable groups.76

In the decade since, government has employed an array of health-related legislation, policies and programs to achieve these goals,77 making some progress. Between 1994 and 2004, over 1,200 new clinics were built (representing 35 percent of primary health care facilities),78 immunization rates of children increased from 63 percent in 1998 to 72 percent in 2002, and the government’s free primary health care policy was extended in 1996 from pregnant and lactating mothers and children under six, to cover primary health care services for everyone in the public health system.79

These measures have not, however, managed to overcome serious and persistent deficiencies in the public health sector, including poor availability of all types of health care personnel at a national level, limited access to critical services such as antenatal care (available at only 50 percent of primary health care facilities),80 negligible access to HIV/AIDS prevention and treatment services,81 and considerable inequities in intra- and inter-provincial health expenditures and health care access.82 Moreover, gross inequities between the public and private systems persist – the public system meets the needs of 80 percent of the population, receives only 40 percent of total health funding and retains only 30 percent of the country’s health personnel.83

*Prima facie* one could argue that these limited improvements in health outcomes reflect the government’s reasonable efforts to progressively realize access to health care. However a deliberate government strategy of curbing health care spending in pursuit of economic goals highlights the naivety of the reasonableness standard which does not inquire into the process of how the overall budget for health care is set. To elaborate, the ANC came into power with a principled commitment to meet basic needs, expressed in the Reconstruction and Development Program (RDP) and adopted in 1994. By 1996, the ANC had abandoned the RDP and in its place had adopted the neoliberal Growth, Employment and Redistribution (GEAR) strategy, which emphasized privatization, deregulation, rationalization of the public sector, and strict economic stringency in social spending.84 As a result, health spending was severely curtailed, and several major tertiary health centers were shut down in the late 1990’s.85

Although health spending increased by ZAR6 billion (US $897 million) from the late nineties to 2005, these allocations have not kept pace with inflation nor with increased demands on the public sector (including an estimated 6.5 million new users since 1996), and the high (and incompletely compensated) cost that HIV/AIDS is exacting on the public system.86 The stagnancy of these allocations raises serious questions about the constitutional reasonableness of government’s health policies given the nexus between progressive realization of access to health care services and the availability of resources.

While the government’s broader macroeconomic policy is certainly designed to improve the country’s wealth, evidence indicates only slight annual per capita increases in the economy.87 These minor economic gains have certainly not improved the well-being of the poorest and most marginalized sectors of the population – indeed there have been increases in poverty and inequality since 1996,88 and a steady rise in unemployment rates.89 Despite some progress in delivering electricity, sanitation and housing,90 there is persistently limited access in many places to adequate social services, including health care.91 Given the negative impact of poverty and inequality on health, these failures exacerbate the country’s poor health conditions, and place additional burdens on the state’s ability to provide health care services for all.

The inadequacy of government efforts on health are borne out by the appalling regression in all major health indicators over the past decade: life expectancy at birth has fallen from 56 years in 1996 to a projected 47 years in 2005, and both infant mortality and maternal mortality have steadily increased since 1998.92 The primary factor in this deterioration is the virtually unchecked growth of the nation’s HIV/AIDS pandemic, the largest in the world.93 From 1.8 million people infected in 1996, infection rates have soared to an estimated 6.29 to 6.57 million people in 2004 – over 14 percent of the population.94 AIDS has become the single biggest cause of death, with over 300,000 people dying per year,95 and an estimated 1.2 million deaths from HIV/AIDS to date.96 In the absence of effective prevention and treatment, more than seven million people will die by 2015, leaving two million children orphaned.97

The public health care system has buckled under the impact of the HIV/AIDS epidemic and its overwhelming illness and death. A 2002 study showed HIV infections in almost half of the patients in public medical and pediatric wards,98 while another study that same year found a 60 percent HIV rate in pediatric admissions in an urban public hospital.99 It is estimated that over time approximately 16 percent of health workers will die of AIDS.100 The epidemic’s negative socioeconomic impact is even broader, and throughout Sub-Saharan Africa HIV/AIDS is deepening household poverty, re-
versing gains in human development, worsening gender inequality and ultimately eroding governments’ ability to maintain essential services. Given the epidemic’s shockingly polycentric impact on health and social welfare, the reasonableness of government policy on health largely depends on how it addresses HIV/AIDS.


Viewed from the perspective of reasonableness, South Africa’s HIV/AIDS policies do not fare well. Indeed the rampant growth of South Africa’s HIV/AIDS prevalence rate suggests a manifest failure of government’s containment strategies. Of course governance failures alone cannot be fingered as the sole cause of an epidemic enabled by multiple social, economic, cultural and epidemiological factors, but there can be little doubt that governance is a critical factor. This is evident in the powerful impact of highest level leadership on effective prevention strategies in Uganda, Senegal and Thailand, and the development of a global consensus about the critical role of leadership in an effective HIV/AIDS response.

South Africa’s President Thabo Mbeki has consistently failed to provide strong leadership to national HIV/AIDS policies, providing direction which has been weak, confusing, counter-productive and even obstructive. Rather than guiding government’s response to HIV/AIDS, he has constrained it. In 1999, the President began to publicly endorse AIDS denialism, whose proponents dispute a causal link between HIV and AIDS. Denialists assert that immune failure attributed to AIDS results instead from the toxicity of antiretroviral treatment (ARV), promiscuous, drug-abusive gay “lifestyles,” and in Africa, from poverty-related malnutrition and illness. These theories, which essentially deny the existence of AIDS pandemics, are discredited in mainstream scientific discourse, and disproved by scientific evidence of the link between HIV and AIDS, and the efficacy of ARV in slashing AIDS-related death rates. The impact of AIDS denialism on governmental policy, and in particular on the government’s approach to AIDS treatment, has been palpable. Cleaving closely to denialism’s claims of the high toxicity of ARV, for many years government effectively refused, delayed and even actively opposed their use in the public sector, despite the desperate need for the drugs, the massive reductions in illness and death that results from their use, and the growing negative impact of the epidemic on health, health care access, and social welfare more broadly.

While the prohibitively high cost of these drugs (US$ 15,000 per annum) justified the government’s initial refusal, by 2002 the cost argument had become increasingly threadbare given considerably reduced drug prices, the government’s legislative authority to itself lower prices further, and multiple new international funding sources for AIDS medicines such as the Global HIV/AIDS Fund. Resources could no longer act as a reasonable justification for inaction, and certainly not when it came to maternal transmission of HIV/AIDS, which accounts for approximately 80-90,000 paediatric infections each year. Nevirapine, a new ARV drug, had been offered to the state free of charge for five years, and strong scientific evidence was emerging of its efficacy in preventing maternal to child transmission (MTCT). Yet the government refused its use citing resource constraints as well as concerns about its toxicity and efficacy.

The lack of principled criteria underpinning this decision, its apparent base in denialism, and its drastic human toll attracted extensive civil society protest and public criticism. The force of these pressures lead the government in 2000 to announce plans for a two year MTCT pilot program using Nevirapine to be run from 18 sites, after which government would assess whether to initiate comprehensive provision. Yet a year later only a few sites were operational and the government had actively blocked the drug’s availability in the public sector until the pilot program was complete.

4.1. Minister of Health and Others v. Treatment Action Campaign (TAC) and Others

In 2001, a number of civil society groups instituted legal action, arguing that prohibiting the use of Nevirapine at public health facilities outside test sites and failing to implement a comprehensive national MTCT program breached section 27 of the Constitution as well as children’s rights to basic health care services in section 28. The government argued that its approach was reasonable given concerns about the safety and efficacy of the drug, resistance to the drug, and the cost of the comprehensive program required. The High Court ruled against the government, which appealed to the Constitutional Court. In its appeal, government repeated these justifications, but also pointedly aimed its appeal at the very act of judicial review of health policy – arguing that the High Court’s order effectively made government policy and therefore violated the doctrine of separation of powers.

The Constitutional Court rejected the government’s arguments vis-à-vis constrained resources. Given the government’s acknowledgement that confining the drug to training sites was unrelated to drug costs, the government’s only costs concerns were for ancillary services related to a comprehensive program (such as testing and counselling, formula feed and monitoring).
However, the Court did not accept that budgetary constraints disallowed a comprehensive program, given evidence lead at the hearing that substantial new allocations had been made for HIV treatment including MTCT.\textsuperscript{111}

The real dilemma from the Court’s perspective was whether it was reasonable to exclude the drug in other public health facilities where testing and counselling were available and the drug was medically required.\textsuperscript{112} The Court dismissed all the government’s arguments in this regard, given the weight of scientific evidence supporting efficacy and safety, the limited threat posed by resistance, and the grave suffering and limited survival prospects of children without the intervention.\textsuperscript{113} The Court stressed that the case was concerned with newborn babies whose lives might be saved by the administration of a simple and cheap intervention, the safety and efficacy of which had been established and which the government itself was providing in pilot sites in every province.\textsuperscript{114} While the Court recognized the need to assess operational challenges and monitor issues relevant to the safety, efficacy and resistance of Nevirapine, it stressed the “pressing need to ensure that where possible loss of life is prevented in the meantime.”\textsuperscript{115}

The operational reasons therefore could not justify their cost in human suffering and loss of life nor could they negate the consequent necessity of urgent action: the interests to be balanced were simply too disproportionate.

The need for urgent action was particularly pressing in the case of HIV/AIDS, which the Court acknowledged was “the greatest threat to public health in our country.”\textsuperscript{116} It is noteworthy that the Court in the T\textsuperscript{4}C decision explicitly indicated to government that the magnitude of the epidemic called for a “concerted, coordinated and co-operative national effort in which government in each of its three spheres and the panoply of resources and skills of civil society are marshaled, inspired and led.”\textsuperscript{117} These are all aspects markedly absent in government’s approach to HIV/AIDS.

The Court reiterated the Grootboom decision’s holdings that programs must not ignore urgent needs nor exclude significant segments of society,\textsuperscript{118} and that state policy in this case affected poor people who could not afford to pay for medical services.\textsuperscript{119} In this case, the Court held that children’s needs are “most urgent” and their inability to have access to Nevirapine profoundly affects their ability to enjoy all rights to which they are entitled. Their rights are “most in peril” as a result of the policy that has been adopted and are most affected by a rigid and inflexible policy that excludes them from having access to Nevirapine.\textsuperscript{120}

The Court firmly rejected government’s arguments that the separation of powers limited it from reviewing health policy and from making anything more than declaratory orders.\textsuperscript{121} The Court strongly affirmed that although the state’s obligations under the Constitution’s socioeconomic rights were qualified, the rights were nonetheless justiciable, and authorized the Court to review health policy.\textsuperscript{122} The Court further emphasized that the Constitution mandated it to grant “appropriate relief”\textsuperscript{123} and make “just and equitable orders,”\textsuperscript{124} and that the constitutionality of mandatory relief and supervisory jurisdiction had been repeatedly confirmed in domestic and foreign jurisprudence.\textsuperscript{125} While the Court declined to exercise supervisory jurisdiction in T\textsuperscript{4}C, its order both declared the government’s responsibility to devise and implement a comprehensive MTCT program, and mandated government to immediately remove restrictions on the drug and make it available in the public sector, provide for the training of counsellors and take reasonable measures to extend testing and counselling facilities throughout the public health sector.

5. The Implications of the Constitutional Right and Jurisprudence

The constitutional right to health care and the Court’s jurisprudence have a number of important implications for health policy, for individual claims to health care and for judicial review. The constitutional right and its interpretation effectively make the national pursuit of universal access to health care an obligation, insulating this decision from the “accidental logics” of political goodwill and historical circumstance that may otherwise exclude a national commitment to equity in health care.\textsuperscript{126} Moreover, pledges to equity fixed in policy alone are easily discarded, as the abandonment of both rationality in HIV/AIDS policy and the RDP in totality illustrate. An entrenched constitutional commitment to equitable health care provides a far more enduring guarantee. The reasonableness standard in particular, mandates a degree of equity in health care by requiring government to devise and implement comprehensive health policies to meet basic health needs and to take particular account of the poor and vulnerable, and of urgent and desperate needs.\textsuperscript{127}

The reasonableness standard also has important implications for the nature of health governance required in fulfilment, which must be open, accountable, comprehensive, coherent and cooperative. This is a powerful guarantee of transparent and responsive health decision-making, and the right and the reasonableness standard should therefore be viewed as serving an important democratic function. Similarly, the right strengthens democracy in the opposite direction, by
and wealthy classes, and where health conditions amongst the poorest populations are in any event deteriorating due to HIV/AIDS. This latter aspect also alerts to a fundamental weakness in the economic argument against increased health allocations, given the recognized necessity of AIDS treatment for economic stabilization and economic development in high prevalence countries.\footnote{131}

However the Court’s approach is not without problems, particularly its refusal to identify substantive content to the right and to recognize or define a minimum core. The Court’s assertions of incapacity are not necessarily convincing given the Committee on Economic, Social and Cultural Rights’ identification of a minimum core to the right to health.\footnote{132} In addition, it is not clear why identifying specific obligations under the constitutional right to health would breach the separation of powers any more than the Court’s identification of specific standards in the reasonableness standard. This is particularly given that the reasonableness standard provides a rough proxy for many of the standards proposed in the international minimum core. For example, the minimum core focuses on ensuring equal access to essential health care facilities, goods and services for the most vulnerable and marginalized social groups.\footnote{133} These are obligations largely implicit in the reasonableness standard, or guaranteed in the Constitution’s other socioeconomic rights.\footnote{134}

While this similarity partially mitigates the Court’s rejection of the core, the reasonableness standard does not sufficiently fill the breach left by insufficient judicial development of the standards that this right imposes.\footnote{135} For example, the reasonableness standard provides little guidance on how government should operationalize the broad constitutional framework it provides for health policy. In particular it fails to provide temporal priorities to guide progressive realization,\footnote{136} and fails to define urgency, desperation, and the key populations of the poor and vulnerable which the reasonableness standard focuses on.\footnote{137} Litigation on discrete issues arising is a poor surrogate for the systemic guidance required to enable effective and timely realization of these standards. For better or worse however, the Court is unlikely to repeal this aspect of its approach to health rights. If the minimum core is to be salvaged, and a comprehensive interpretation of the substantive content of section 27 is to be achieved, it will have to be done outside the Court, perhaps through the development of a nationally appropriate “General Comment 14” on the constitutional right to access health care services.\footnote{138}

The Court’s refusal to adopt the core therefore should be seen as evidence of a cautious approach to adjudication which does not entirely refute the idea that enforcing socioeconomic rights may breach the doctrine

---

Lisa Forman
of separation of powers. Yet the Court’s explicit desire to adopt a “restrained” role to adjudication seems less related to any breach the core could cause, than to signal judicial restraint to government in order to gain the trust of the legislature and executive, thereby bolstering its own institutional legitimacy. The Court’s restrained approach to enforcement therefore reflects its desire to appear cautious in developing its own powers within South Africa’s growing constitutional democracy, rather than any inherently unenforceable aspects of health rights themselves. This cautious approach is somewhat counter-balanced however by the Court’s firm assertion of constitutional authority to review socio-economic policies and to make creative and far-reaching remedial orders if necessary, and this alone significantly advances broader understanding of the possibilities of socioeconomic rights adjudication, even though the Court’s approach remains rooted in the old paradigm attending to socio-economic rights.

Given transnational judicial referencing and widespread respect for the Constitutional Court, the South African experience therefore may hold powerful didactic value for how judges themselves perceive the possibilities of enforcing health rights, and the poverty of the old paradigm as a bar to adjudication. Given transnational judicial referencing and widespread respect for the Constitutional Court, the South African experience therefore may hold powerful didactic value for how judges themselves perceive the possibilities of enforcing health rights, and the poverty of the old paradigm as a bar to adjudication.

Conclusion
The South African experience suggests that the influence of health rights on health policy is strongly contingent on a number of factors, including whether health rights are legally entrenched, politically recognized, judicially enabled, and socially claimed. Where several of these factors converge as they have in South Africa, the capacity for these rights to fulfil their transformative potential comes close to being realized. Nonetheless, the narrow scope of the constitutional health right and the necessarily case-specific intervention of the Court suggest that the power to assure good public health lies squarely with the government, which alone has the capacity to ensure adequate sustained efforts regarding public health care, and to address the
underlying determinants of health such as water, electricity, food and sanitation. Courts can nonetheless play an important role in enabling discrete instances of access to health care services, and in remedying gross policy defects that may have very broad impacts on health and health care access. Finally, by ensuring that the government comply with fundamental human rights values and constitutional protections in its health policies, rather than derogating from democratic accountability, enforcing health rights holds the potential to increase it.

Acknowledgements
This article is based on research funded by a Canadian Institute for Health Research Health Law and Policy Training Fellowship. The author thanks Colleen Flood, Juliette Nicolet and Tracy Cohen for comments on earlier versions of this article.

References
1. See for instance, Dandridge v. Williams 397 U.S. 471 (1970), where the U.S. Supreme Court refused to recognize the way that social grants were formulated or administered, reasoning that “[t]he intractable economic, social and even philosophical problems presented by public welfare assistance program are not the business of this Court.”
2. See for example, Andrews v. Law Society of British Columbia (1989) 1 S.C.R 123, at paragraph 194, where Justice La Forest states: “much economic and social policy making is simply beyond the institutional competence of the courts: their role is to protect against incursions on fundamental values, not to second guess policy decisions.”
3. See for example, Gosselin v. Quebec (Attorney General), 2002 4 S.C.R. 429, where the Canadian Supreme Court refused to recognize that the right to security of the body places positive obligations on the government to provide social welfare.
7. There are examples of both direct judicial enforcement of health rights, and judicial protection of health through civil rights protections. For example in Viceconti v. Ministry of Health and Social Welfare (Argentina, Poder Judicial de la Nación, Causa no. 31.777/96, June 2, 1998), international human rights treaties on health were recognized, and government’s positive obligations to provide health care were confirmed; and in Cruz Bermudez et al. v. Ministerio de Sanidad y Asistencia Social (Supreme Court of Justice of Venezuela, Case No. 15.789, Decision No. 916 July 15, 1999), the court found that constitutional rights to health and life required the Ministry of Health to provide antiretroviral medicines; develop national treatment policies and programs; and reallocate budget necessary to carry out the Court’s decision. For indirect protection of health, see for example, Paschim Banga Khet Mazdoor Samity v. State of West Bengal (1996) 6 S.C.C. 37, where the denial of emergency medical treatment for serious head injuries was deemed to be a breach of the right to life contained in the Indian Constitution; and Eldridge v. British Columbia (Attorney General) 1972-3 S.C.R. 624, where the Canadian Supreme Court ordered the provision of health care benefits under the constitutional right to equality and non-discrimination.
10. Id, Preamble.
11. Id, sections 1(a), 7(1), 36(1) and 39(3)(a).
12. Id, Preamble.
14. These rights include section 27 (food, health care, water and social security), 26 (housing), 29 (education), and 28 (children’s rights to basic social amenities).
15. Under apartheid, approximately 87 percent of South African land was reserved for white ownership, and what limited black ownership remained was effectively eroded through forced removals. Skilled jobs and economic opportunities were predominantly limited for whites.
16. South Africa is ranked 119th on the United Nations Development Programme’s (UNDP) human development index (in the lower two thirds of countries classified as having medium human development). However it has a gross domestic product per capita (US $10,070, 00) comparable to at least 60 percent of countries classified as having high human development. See UNDP, Human Development Report 2004 (New York: UNDP, 2004): at 138-139, 7. K. Steyn and M. Schneider, “Overview on Poverty in South Africa” in D. Bradshaw and K. Steyn, eds., Poverty and Chronic Diseases in South Africa: Technical Report 2001 (Tygerberg: Medical Research Council, 2001): 1-14, at 6. S.A. records one of the highest Gini Coefficient indexes in the world, measured in 2003 at 59.3. The Gini coefficient is used to compare inequality among nations. It measures the distribution of income (or consumption) among individuals or households within a country between 0 (perfect
equality) and 100 (perfect inequality).


19. See Steyn and Schneider, supra note 17, at 12.


21. Constitution, supra note 9, at sections 28(1)(c), 35(2)(e) and section 24(a), respectively.

22. See for example, WHO Constitution and ICESCR, supra note 4.


25. There is however an important proviso to this limited scope, since certain health care services (such as AIDS medicines in high prevalence countries) may hold far broader public health benefits. Moreover, given the correlation between poverty and health, the realization of the constitutional provisions regarding basic needs could also contribute to improved population health.

26. Constitution, supra note 9, at section 7(2).

27. The duty to respect imposes a negative obligation to desist from interfering with people’s enjoyment of rights; the duty to protect requires the state to prevent third party interference with people’s rights; and the duty to promote and fulfills the state’s positive obligation to realize access. The notion of a typology of rights is widely acknowledged to have been developed by Henry Shue, see H. Shue, Basic Rights, Subsistence, Affluence and US Foreign Policy (Princeton, NJ: Princeton University Press, 1980). It is applied in the context of health in General Comment 14, supra note 5.


29. Id.

30. Id., at paragraph 78.

31. These concerns are apparent in Justice Albie Sachs’s ex curia assertion, prior to the Court’s seminal decision in Grootboom, that judicial lack of expertise regarding the technical complexities of socioeconomic policy required “corresponding judicial modesty.” In Sachs’s eyes, judges could not “be philosopher kings and queens who go around telling government how to function.” Where Sachs saw an appropriate judicial function, was when situations of socioeconomic policy required “corresponding judicial modesty.” Instead Soobramoney argued that the state could make additional funds available to the provincial hospital, and that it was obliged to do so under section 27. Soobramoney, supra note 33, at paragraph 23. See also Chaskalson, “Moral Transformation,” supra note 8, at 603.

32. Soobramoney, supra note 33, at paragraphs 24-25.

33. See id. at paragraph 42.

40. Id., at paragraph 29 states that “a court would be slow to interfere with rational decisions taken in good faith by the political organs and medical authorities whose responsibility it is to deal with such matters.” Rationality is a low standard of review, requiring only that the government purpose is legitimate, and that there is a rational and not arbitrary connection between the law and the government purpose. See also Khosa, supra note 32, at paragraph 67.

41. Grootboom, supra note 34, at paragraphs 24 and 44.

42. TAC, supra note 35, at paragraphs 35 and 36.

43. Grootboom, supra note 34, at paragraph 83.

44. Id., at paragraphs 38, 40, 42 and 43.

45. Id., at paragraph 44, cited with approval in TAC, supra note 35, at paragraph 68.

47. Grootboom, supra note 34, at paragraphs 35 and 43, the latter paragraph cited in approval in TAC, supra note 35, at paragraph 68.

48. Grootboom, supra note 34, at paragraph 44.

49. Id., at paragraph 44.

50. Id., at paragraph 45.

51. Id., at paragraphs 40 and 43; cited with approval in TAC, supra note 35, at paragraph 68.

52. Grootboom, supra note 34, at paragraph 41.

53. Id., at paragraph 45.


56. TAC, supra note 35, at paragraph 131.

57. Id., at paragraph 32.

58. Grootboom, supra note 34, at paragraph 46.

59. Id., at paragraph 68.

60. Soobramoney, supra note 33, at paragraphs 8-11; TAC, supra note 35, at paragraphs 31 and 34; Khosa, supra note 32, at paragraph 43.

61. Soobramoney, supra note 33, at paragraphs 8, 28 and 31.

62. Id., at paragraph 53.

63. Id., at paragraph 54.

64. Id., at paragraph 32.


67. The minimum core is developed in relation to health in General Comment 14, supra note 5, at paragraph 42.

68. Grootboom, supra note 34, at paragraph 33.

69. TAC, supra note 35, at paragraphs 37 and 38.

70. Id., at paragraph 38.

71. Id.

72. Id., at paragraph 36.

74. For more comprehensive assessments along these lines see K. Pillay, “Tracking South Africa’s Progress on Health Care Rights: Are We Any Closer to Achieving the Goal?” Law Democracy and Development (2003); 7; and the annual socio-economic rights reports by the South African Human Rights Commission, see <http://www.sahrc.org.za> (last visited October 4, 2005).


81. G. Reagon et al, supra note 78, at ix. Despite major advances in national AIDS policies on treatment for mother to child transmission of HIV, and a national ARV plan discussed below, access remains greatly limited.


90. South African Government Communications, supra note 79.


100. Shisana et al, supra note 98, at xiv.


There can be little doubt that political protest and the upcom-

There is a natural overlap between the reasonableness standard's

Concerns about the polycentric nature of adjudication (partic-

The minimum core focuses on meeting the needs of the poor and vulnerable, and the

The task of defining a nationally appropriate minimum core is

The only specific minimum core obligation not explicit in the

The minimum core is drawn from P. Braveman and S. Gruskin, “Defining Eq-

focus in health equity on addressing systematic health differences

The latter definition is drawn from P. Braveman and S. Gruskin, “Defining Eq-

The inter-relationship between them mandated by the Constitu-

The delicate balance called for to give effect to this separation re-

The independence of the judiciary and the separation of powers are foundational principles of constitutionalism. The

Health Law 2005: An Agenda

Peter D. Jacobson

In 2004, the journal Health Matrix published a very interesting symposium volume titled “The Field of Health Law: Its Past and Future.” As the title implies, the various commentators took both a retrospective and a prospective look at past trends and future prospects in health law. Some, including Clark Havighurst, Skip Rosoff, and Walter Wadlington, wrote thoughtful essays on the development of health law over time and the implications of those trends. Others, including Rob Schwartz, Jim Blumstein, Rand Rosenblatt, and Mark Hall and Carl Schneider, wrote equally thoughtful essays that reflected on the past but focused more on future directions and prospects. And one, Ken Wing, wrote a semi-dyspeptic essay debunking the entire field of health law.

Taken together, these essays present a comprehensive view of how health law has developed so far and where its future might lie. Four themes emerge from the collected writings. First, there is considerable agreement on how and why health law has developed, but little agreement on where it is headed. Second, there is considerable concern as to health law’s place as a compelling discipline within the law school curriculum. Many of the essays might therefore be read as self-reflective attempts to assess the field and the authors’ collective contributions to it. Third, many of the articles take a position on the desirability of markets versus professional norms or social justice in defining the role of law in the delivery of medical care.

Fourth, the volume implicitly (and explicitly in Ken Wing’s contribution) raises questions about the field’s current dynamism. In this context, consider the following observation from Mark Hall and Carl Schneider: “We suspect there is no grand organizing principle for medical law because there cannot be. Medical law deals with medical activities in too many settings and must borrow from too many areas of law....We propose an analytical framework that views health care law as a law of relational webs rather than a law of transactions.” Somewhat more pithily, Rob Schwartz suggests that we just “follow the money!” Both observations are intriguing and offer alternative ways of viewing the field’s current status.

Yet despite the ruminative qualities in the Health Matrix volume, in many ways medical liability continues to dominate health law. At least in the policy sphere, medical malpractice reform commands so much attention that it effectively eclipses other equally important areas. And the number of health law scholarly articles regarding medical liability continues apace. Just

Peter D. Jacobson, J.D., M.P.H., is Professor of Health Law and Policy at the University of Michigan School of Public Health, and Director, Center for Law, Ethics, and Health.
When I think there is nothing left to say on the subject, several new articles appear. Medical liability may not be the essence of the health law endeavor to scholars (as the contents of any health law casebook will demonstrate), but I suspect that much of the public at large views malpractice as the core of health law.

In this article, I will take a more pragmatic look at health law in 2005 than in the *Health Matrix* volume, focusing on some areas that have not received as much attention as they deserve. Assuming that the current state and federal tort reform efforts succeed in reducing the current medical liability crisis, attention will shift to other pertinent health law issues. What follows is an agenda for post-tort reform health law. In the second part, I address what I consider to be the most significant issues that directly affect health law doctrine in the short-term. In the third part, I examine some issues that are largely about the intangible and long-term aspects of the physician-attorney and law-medicine interaction. In the fourth part, I will offer some observations about trends in legal doctrine involving the health care delivery system.

**Developing Sound Legal Doctrine**

Despite the tendency of medical liability to occupy the existing policy and doctrinal spheres, there’s actually a robust, if less visible, debate about the direction of legal doctrine in health care delivery. In this section, I want to focus on two broad areas of interest – competition policy and patient safety. In my view, these represent the most immediate concerns facing health law, particularly in determining the primacy of markets versus deference to medical professionals in shaping health law doctrine and health policy.

**Competition Policy**

Competition policy, the role of the government in facilitating a free-market health care system, lies at the heart of how the health care industry is organized and raises a set of interesting and challenging issues at the epicenter of health law. Each of these would warrant a separate presentation to cover adequately, so I’ll just outline them for now. The core questions animating competition policy are: Who owns the health care enterprise? To what extent should health care be a market-driven industry, accountable primarily to consumers and shareholders? Or, should government intervene either to correct market failures or to alleviate the inevitable distributional consequences of a market-driven system? How does the framing of the question affect the development of legal doctrine?

**Governance**

Unlike most competitive markets, where the dominant organizational form is a for-profit corporation, the health care industry is comprised of a mix of for-profit and nonprofit organizational forms. Health care administrators often use the phrase “no margin, no mission,” to suggest that meeting the charitable mission is impossible unless the organization generates sufficient revenue to sustain its operations. Thus, nonprofit health care administrators must balance efforts to maintain adequate margins without compromising the charitable mission. In the managed care era, any given health system might integrate both for-profit and nonprofit forms within its overall structure, and might engage in joint ventures based on a mix of organizational strategies. But these strategies raise complicated legal issues involving antitrust, tax exemption, and regulatory compliance. They also raise questions of governance and fiduciary duties.

Fiduciary duties are rooted in notions of professional norms, but operate within the context of a competitive market. The primary governance fiduciary duties, of care and obligation to the institution, are rooted in the recognition that trustees have an obligation to operate the facility for the community’s benefit, not for their own pecuniary interests. Nonprofit fiduciaries must therefore balance their obligations to the community with the realities of changing markets (i.e., balancing equity and efficiency). At a minimum, fiduciaries must consider intangibles that the market paradigm may not easily value, such as meeting the IRS’s ill-defined community benefit test. Fiduciaries have considerable flexibility to operate within the business judgment rationale, but must be careful to protect the organization’s assets while still meeting the institution’s community benefit obligations.

To say the least, governance is anything but a “sexy” area of the law. In my view, however, it is of fundamental importance in health care because of the competing pressures of mission versus margin. A major problem
is that the law of fiduciary duty is underdeveloped and inadequate to meet the competing needs of executives running a modern health care organization. Health law scholars are just beginning to address these issues. For example, Tim Greaney and Kathleen Boozang argue that fiduciary law is “muddled and too permissive in its oversight” but that the competing doctrine of charitable trust law is “doctrinally inapposite and pragmatically unsuited to govern business conduct in the contemporary health care market.”

Instead, they advocate a normative standard called “mission primacy,” which they define as “a doctrinal recognition that the nonprofit corporation’s articulated charitable mission is its central objective. Mission primacy...would extend the concept of the duty of obedience to underscore directors’ core responsibilities as stewards of a nonprofit enterprise to advance its public purpose.” As with current fiduciary duty doctrine, their approach would accord considerable deference to trustees, with a more exacting requirement of considering the charitable mission in all decisions.

At first reading, I fully agreed with the central tenet espoused, but found mission primacy to be too vague for useful doctrinal development. On reflection, though, I think there’s considerable merit to this approach because it offers an overarching theory that can then be more precisely defined through subsequent common law development. I would nonetheless amend their standard to place the charitable mission in a paramount position, requiring compelling evidence to permit any deviation or departure from the mission.

A good example of how the mission primacy standard might work (even though it would not have changed the result) is the case of In the Matter of Manhattan Eye, Ear & Throat Hospital v. Spitzer, or MEETH for short. In MEETH, the board of trustees of a nonprofit acute care hospital specializing in ophthalmology, otolaryngology, and plastic surgery decided that the facility could no longer be competitive and needed to sell its assets. To make a long story short, the board hired a consulting firm, which concluded that the business had no value and that the underlying real estate was the only valuable asset. Ultimately, the board voted to monetize the value of the real estate and make no effort to preserve MEETH’s charitable health care mission. The New York State Attorney General challenged the board’s decision essentially to close the hospital and sell it for the value of the real estate. Based on its analysis of the fiduciary standard of duty of obedience, the court ruled that the sale violated New York law, stating that:

...While it may be appropriate, in certain cases, to solve financial difficulties by eliminating the organization’s mission by selling its assets and then undertaking a new mission,...the duty of obedience...mandates that a board, in the first instance, seek to preserve its original mission. Embarkation upon a course of conduct which turns it away from the charity’s central and well-understood mission should be a carefully chosen option of last resort. Otherwise, a board facing difficult financial straits might find sale of its assets, and “reprioritization” of its mission, to be an attractive option, rather than taking all reasonable efforts to preserve the mission which has been the object of its stewardship.

Clearly, in my view, the court reached the right result. The question is whether the mission primacy standard would provide a stronger doctrinal rationale than the court’s reliance on the fiduciary duty of obedience. I believe that the answer is yes. The two doctrines have different starting points and the mission primacy standard places the burden on the trustees to demonstrate why mission primacy should not prevail. The duty of obedience test effectively places margin and mission on an equivalent plane. Trustees have greater flexibility to argue that business judgment requires the sale or conversion of the assets to non-charitable purposes. In contrast, the mission primacy standard arguably forces the parties to start with the charitable mission as the preferred option and should only be rejected if the trustees present compelling evidence that continuing the mission is impossible.

But to demonstrate how subjective and potentially difficult the “mission primacy” standard might be doctrinally, consider that my analysis of the application of the standard to the MEETH case differs from Professors Greaney and Boozang. They note that “[i]n this regard, mission primacy would likely have required a less categorical evaluation of purpose in MEETH,” whereas I view the result as being more categorical (or at least no less so). Nevertheless, mission primacy is a promising start for rethinking legal doctrine with regard to governing the nonprofit health care enterprise.

Professor Jill Horwitz has also been examining governance issues and proposes that all hospitals should be governed by a duty of integrity, defined as a moral constraint on organizational behavior. The duty of integrity would operate as follows:

A for-profit hospital can decide which services to provide based on the profitability of the service, as long as that decision does not violate the ethical responsibilities of the hospital as a healthcare provider. This means that if the hospital invests in a service because it is profitable it must invest in related, unprofitable services....Not-for-profit hospitals cannot decide to offer services solely to gener-
ate profits, but could do so to subsidize other services that they could not otherwise offer. Under this theory, a not-for-profit hospital could market special services to particularly high-paying clientele to subsidize unprofitable services. 17

While the content of the duty would differ across organizational forms, the duty would impose contractual obligations for health care and health care justice. Allowing for differences in moral obligations according to form (that for-profits can pursue profits per se, unlike nonprofits), the duty of integrity would require all facilities to benefit the public. Whether this is a workable framework remains to be determined (though I have no doubt that the board in MEETH would not have satisfied the duty of integrity), but it is certainly a useful complement to the Greaney and Boozang approach.

**Tax Policy**

A set of issues closely related to governance considerations is the increasingly scrutinized tax exempt organizational form. To the extent that the tax exempt form continues to be integral to the organization of health care delivery, the debate over governance remains salient. But the question we might ask is whether the tax exempt form is obsolete in an increasingly competitive industry? If so, the governance debate will shift dramatically toward defining and ensuring a level of community benefit that for-profits should provide.

The battle over the tax exempt form is two-fold: whether markets or a social justice paradigm should prevail; and whether nonprofits actually justify their tax exempt status. Supporters of the competitive market paradigm argue that for-profits provide relatively equal amounts of free care, contribute tax revenue to the community, and provide more efficient health care delivery. 18 More cynically, market proponents might add that the social justice model of publicly funded charity hospitals has failed because the public refuses to invest resources into facility modernization. Accordingly, only the private sector can “save” charity hospitals by converting them to for-profit status. 19

Market proponents argue that since it is more efficient to encourage institutional integration for providing health care, the resulting efficiencies will generate additional tax revenue to provide adequate amounts of uncompensated care. Proponents of social justice and equity in the distribution of health care are likely to counter that the value of community benefits nonprofits provide, especially when those services would be otherwise unavailable, should outweigh efficiency gains. For efficiency to outweigh equity, market proponents should have the burden of demonstrating that a more efficient delivery system will accommodate those unable to pay.

Regardless of how that conceptual debate proceeds, the IRS and local municipalities are increasingly scrutinizing tax exempt facilities regarding their compliance with community benefit expectations. Given that there are few, if any, operational differences between for-profit and nonprofit health care facilities (known as convergence – the differences are in how the revenues are distributed), what is the justification for continuing the tax exempt form? One of the policy goals supporting tax exempt facilities is to provide health care for those unable to pay. Historically, the presumption has been that for-profits will shirk their community responsibilities. Yet several scholars dispute that characterization. Since many public hospitals face severe financial stress, they may no longer be able to provide levels of free care as in the past. And as cash-strapped state and local governments seek new revenue sources, they will be tempted to scrutinize closely tax exempt health care facilities to ensure that they are providing commensurate community benefits in return for the tax exemption. 20

In a competitive environment, nonprofits are forced to cut back on the unprofitable services they provide. 21 This has led to a series of unsuccessful lawsuits claiming that nonprofits have overcharged uninsured patients. Further, competition has compelled nonprofits to seek revenue generating sources, often through joint ventures with for-profit entities. Without going into the legal intricacies of these deals (i.e., Revenue Ruling 98-15), the competition policy question is whether these arrangements erode the charitable mission, in part because insiders may benefit in violation of the IRS’s private benefit/private inurement rules (prohibiting nonprofit insiders from more than incidental benefits from a transaction).

Professor John Colombo is highly critical of the IRS’s analysis of private inurement/private benefit because it compromises legitimate revenue-generating ventures with an overly broad interpretation of private benefit. Colombo argues that the IRS has expanded the definition of private benefit such that the exemption could be revoked in any given transaction if the private benefit outweighs the public benefit (regardless of the financial enhancement to the exempt purpose). 22 In arguing against the IRS’s private benefit/private inurement analysis, Colombo states colorfully: “What, exactly, does it mean to say that a joint venture impermissibly serves the ‘private interests’ of a for-profit investor? For cryin’ out loud, a joint venture between a nonprofit organization and for-profit investors is supposed to serve the investors’ ‘private interests’ (e.g., make them money) – otherwise, they wouldn’t be part of the deal. And so what if it does?” 23
Antitrust

Any analysis of competition policy should really start with developments in antitrust doctrine. Antitrust enforcement lies at the heart of determining the extent to which health care will be dominated by professionalism or by market competition. But for all the sturm und drang over antitrust, it appears as though antitrust enforcement has not generated the changes its proponents had anticipated. For example, the FTC has lost most of its challenges to mergers, and physicians have notably failed in using antitrust litigation to challenge staff privileges decisions. Nevertheless, the application of antitrust law has unmistakably altered the competitive environment in at least two ways: by not protecting other competitors from managed care, and by supporting managed care’s cost containment initiatives at the expense of physician dominance.

In some ways, the most important doctrinal developments will be in how courts integrate quality of care into antitrust analysis. For health care antitrust cases, developing adequate doctrine to incorporate quality of care considerations is essential. For instance, a range of cases, including joint ventures and mergers, will be resolved very differently depending on how potential quality of care gains will be weighed against the anticompetitive aspects of the arrangement. As Peter Hammer and Bill Sage have noted, courts have generally not consistently applied or defined non-price considerations (i.e., quality of care) in health care antitrust analyses. Take, for instance, the following quote from Blue Cross & Blue Shield United of Wisconsin v. Marshfield Clinic:

...when dealing with a heterogeneous product or service, such as the full range of medical care, a reasonable finder of fact cannot infer monopoly power just from higher prices – the difference may reflect a higher quality more costly to provide – and it is always treacherous to try to infer monopoly power from a high rate of return....One HMO may charge higher prices than other HMOs (and Security does charge higher prices) not because it has a monopoly but because it is offering better service than the other HMOs in its market. Compare itself stresses the quality of the Marshfield Clinic’s doctors, as part of its argument that it cannot succeed unless the Clinic is forced to join it. Generally you must pay more for higher quality.

Even if conceptually accurate, the opinion does not define what quality of care means or how it should be weighed against price considerations.

To be fair, it is not surprising that judges might be reluctant to define quality of care, since health services researchers disagree about how to define and measure quality. And while antitrust law must reconcile trade-offs between price and quality, physicians have historically opposed intrusions into their clinical domain. Because it is increasingly likely that quality of care will be used to justify arrangements that would otherwise be suspect as anticompetitive, this is fertile territory for health law scholars.

In particular, an interesting doctrinal issue in cases involving quality of care is whether antitrust doctrine will reject professional norms in favor of market mechanisms. To Professor Havighurst’s dismay, the Supreme Court held in California Dental Association v. Federal Trade Commission, that the California Dental Association’s restrictions on advertising did not constitute a naked restraint on trade. The Court rejected the FTC’s quick-look analysis under the rule of reason and held that the agency needed a more detailed review to determine whether the restrictions might have procompetitive effects. Professor Havighurst criticized the opinion as substantially expanding potential defenses based on professional concerns. "Indeed the CDA opinion is the most explicit one to date that embraces the notion that professional self-regulation may directly restrain competition yet still be open to justification by demonstrating the existence of a theoretical market failure that the restraint in question may help to overcome." If the California Dental reasoning were to be extended to quality of care defenses, professional norms could well emerge as a strong defense to antitrust challenges.

Fraud and Abuse Regulations

A fundamental aspect of competition policy is the role of regulatory oversight. At this point, whether the current regulatory environment is burdensome or beneficial (or creates situations where compliance with one regime is inconsistent with another) has not been adequately explored. Among the various regulatory programs, the fraud and abuse regime is perhaps the most significant one to consider because of the high compliance costs and the potential impediments to beneficial and efficient arrangements between physicians and
health systems. A good example is the concept of gain-sharing, where physicians share in costs saved by MCOs. One branch of the government, Congress, encourages such innovations, while another, the Office of the Inspector General (OIG – DHHS), initially condemned the concept as violating the fraud and abuse statutes. Recently, OIG has backed off a bit to allow these arrangements under certain circumstances, but has yet to characterize gainsharing plans it would find acceptable.

Surprisingly, health law scholars have paid only limited attention to this area. Some scholars, including Jim Blumstein and David Hyman, have taken a market approach to argue that managed care would alleviate the need for governmental oversight of fraud and abuse. In contrast, Joan Krause has argued that fraud and abuse may be viewed as one aspect of the government’s quality of care oversight mechanism. Yet developing an alternative regulatory framework that would prevent egregious examples of fraud and abuse without the onerous structure now in place has not been developed. While I recognize that health law scholars might not find this an attractive area because of its lack of theoretical content, it is vitally important to health care practitioners and health care delivery.

**Summary**

The various aspects of competition policy covered above offer considerable potential for empirical and conceptual health law scholarship. What they may lack in theoretical excitement is more than offset by the impact health law scholars can have in shaping competition policy doctrine and influencing the structure and delivery of health care.

**Patient Safety**

Patient safety concerns have become a dominant feature of the health policy/regulation debate. A key point of intersection between the patient safety movement and medical liability is whether to report adverse medical events and, if so, whether the confidentiality of information can be protected. Firms and physicians are not likely to support reporting requirements that would expose them to liability. Thus, the question is whether retaining the current liability system, which focuses on an individual patient’s care, is superior to cooperative alternatives more broadly concerned with systems improvements to protect patient populations.

Most of the patient safety discussion has appropriately focused on systems improvements without adequately considering the management/governance failures that contribute to systemic deficiencies in quality of care. As a result, some of the previous discussion of governance would be applicable to patient safety. In addition, changes in legal doctrine could enhance and encourage the adoption of patient safety systems. For example, accrediting bodies have developed sentinel event reports to encourage facilities to report errors and devise strategies for system-wide patient safety improvements. These are voluntary reports that accreditors will keep confidential (though the reports can probably be obtained during litigation).

Professor Bryan Liang has been a proponent of error disclosure without recrimination, and I agree with his analyses about the need to focus on preventing future errors rather than placing blame on an individual physician or facility. This is not to suggest that liability litigation for substandard care should be abandoned. To the contrary, permitting blame-free error disclosure puts the emphasis on deterrence without undermining other desirable tort system goals, such as compensation. Plaintiffs’ attorneys will simply need to gather evidence of substandard care from alternative sources rather than being handed what amounts to an admission of error. By analogy, we protect peer review information even though it may be damaging to a physician whose staff privileges may be denied. Likewise, we mandate child abuse reports and provide physicians with immunity even if the allegations prove to be false.

This general approach has now been codified in the recently enacted Patient Safety and Quality Improvement Act of 2005. Through the establishment of patient safety organizations (PSOs), the Act creates a system for confidential reporting of adverse medical events. Health law scholarship can contribute to the ways in which the Act is implemented by developing regulatory strategies to ensure that PSOs achieve the Act’s goal of balancing patient safety with accountability for medical error.

The area of patient safety has received appropriate attention from health law scholars. Since patient safety is likely to be on the policy agenda for the near future, further scholarly contributions and attention are certainly warranted.

**Addressing Broader Issues in Law and Medicine**

Beyond the purely doctrinal world discussed above, there are some important intangible issues shaping the law-medicine interaction that health law scholarship could productively illuminate. I use the word intangible because the doctrinal implications are not obvious for two of the three areas I will consider, and because these issues are as much contextual as they are subject to doctrinal development. Understanding them in context helps make sense of where the field is at present and identifies some aspects that require additional thinking.
Reducing Health Care Inequalities –
The Absence of Judicial Leadership
If health care delivery and law were driven by a social justice model (as discussed below) as opposed to a market approach, reducing health care inequalities might be of lesser import for health law scholarship. As it stands, market competition drives health care delivery. Despite claims that the market will correct inequalities, there is little evidence showing that market arrangements have effectively addressed inherent inequalities, so it seems appropriate to consider possible judicial interventions.

Since courts are unlikely to impose a right to health care, the primary opportunity to address inequalities is through statutory interpretation. When interpreting legislation, courts are often presented with opportunities to develop doctrine that would reduce health inequalities. Among the many possible examples, four particular instances show the judiciary’s opportunities to confront health inequalities. Even though the results have not been promising for those looking to the courts for leadership, it is important for health law scholars to continue pressing for doctrinal changes that would reduce these disparities.38

Medicare and Medicaid
By their very nature, the Medicare and Medicaid programs are characterized by economic inequalities. Medicare beneficiaries, for example, must purchase supplemental insurance for certain services. Medicaid benefits vary across states and are much less generous than Medicare, offering opportunities for judicial intervention. As Colleen Grogan and Eric Patashnik note, the Medicaid program is rife with equity concerns, particularly with regard to differential benefits for poor families and the elderly, blind, and disabled.39 Indeed, they argue that Medicaid has served to exacerbate existing health inequalities. Repeatedly, however, the courts have refused to redress inequalities in government health care programs.

In a landmark case, Alexander v. Choate,40 Medicaid beneficiaries challenged the Tennessee Medicaid program’s fourteen-day limitation on annual inpatient days as a violation of Section 504 of the Rehabilitation Act of 1973. The beneficiaries argued that the limitation disproportionately affected persons with disabilities who would require more inpatient care. In rejecting the challenge, the Supreme Court suggested that Congress did not intend to require the state to provide specific levels of care and that the state retained considerable discretion in how to structure its program. Significantly, the Court signaled its unwillingness to address any resulting inequalities by stating that:

…to require that the sort of broad-based redistributive decision...always be made in the way most favorable, or least disadvantageous, to the handicapped...would be to impose a virtually unworkable requirement on state Medicaid administrators.

Alexander v. Choate characterizes the prevailing judicial attitude toward addressing deficiencies in state and federal programs. It is hard to imagine a more explicit statement opposing an expansive role in redressing social inequalities. Since this case, the courts have rarely taken affirmative steps to address inequalities in governmental health care programs, despite increasing evidence of widening racial and socioeconomic disparities in Medicare and Medicaid. For instance, in a case challenging Tennessee’s planned procedures for reducing enrollment in TennCare, the Sixth Circuit Court of Appeals ruled that the state’s plan does not violate either Medicaid laws or regulations or constitutional due process requirements.41 Instead of addressing the potential inequalities resulting from the state’s plan, the court ruled that the changes were administrative in nature and therefore did not generate due process requirements.

EMTALA
Congress enacted the Emergency Medical Treatment and Active Labor Act of 1986 (EMTALA) in response to concerns about “patient dumping.”42 This occurs when patients who are unable to pay are refused emergency medical treatment or are transferred from one hospital to another before their condition has been diagnosed and stabilized. Underlying the EMTALA statute is the assumption that people have a right to at
least basic emergency medical attention regardless of insurance status or ability to pay. EMTALA requires that a patient be diagnosed and stabilized before being transferred to another facility. Most states have more stringent requirements for emergency departments to provide care of last resort.

The original legislative intent of EMTALA was to reduce discriminatory practices in emergency rooms and prohibit socioeconomic inequalities in access to emergency medical treatment. From its inception, the legislative goals of EMTALA were not well defined, allowing considerable room for judicial interpretation. The Supreme Court first addressed EMTALA in Roberts v. Galen of Virginia, holding that a plaintiff does not need to prove that the failure to provide emergency treatment resulted from an improper motive. But the Court noted that EMTALA’s requirements are limited in scope to stabilizing and screening a patient. EMTALA does not establish a federal standard of emergency medical care and the courts have refused to broaden EMTALA to address inequalities in access to basic medical services. As a general proposition, courts have deferred to Congress for any such expansion. Under current law, where no emergency situation (or other exception) exists, physicians retain the authority to refuse treatment to any individual without facing legal liability.

To be sure, some opinions have expanded EMTALA’s scope. For example, courts have held that EMTALA applies to all patients in the hospital who have an emergency condition, not just those who arrive for treatment in the emergency room. Courts have also applied EMTALA to all parts of a hospital or health care system. Nonetheless, most judicial interpretations of EMTALA have been restrictive, holding that the courts should not address the efficacy or appropriateness of medical procedures performed under the statute. Most courts have ruled that EMTALA ensures only that the hospital’s protocol is uniformly followed regardless of a patient’s ability to pay.

**Hill-Burton**

The 1947 Hill-Burton Act was passed in response to a perceived national shortage of hospitals, and provided federal monies for the unprecedented expansion of the U.S. hospital sector. The primary intent of the Hill-Burton Act was to provide federal financial assistance to states to “provide for adequate hospitals and other facilities” and “to furnish needed services for persons unable to pay.” To this end, the Act required that the facilities receiving assistance must: 1) make the facility available to all persons residing in the territorial area, and 2) provide a “reasonable volume of services” to persons unable to pay. Unfortunately, these two stipulations – the “community service obligation” and the “reasonable volume” (or “uncompensated care”) assurances – were not well defined and initially not widely enforced.

The Hill-Burton obligations were at best a vague and limited legislative attempt to address economic and social inequalities in the provision of health care services. From 1946 to 1976, the Hill-Burton program supported the construction of forty percent of hospital beds in the United States. But hospitals receiving Hill-Burton funds largely ignored the community service and reasonable volume requirements. These requirements were not enforced until the late 1970s, when federally funded legal service lawyers persuaded the federal courts to enforce the Hill-Burton statute. During this time, advocacy groups for indigent patients successfully sought to use Hill-Burton to establish a right to health care for the indigent.

In 1979, primarily in response to Hill-Burton litigation, the Department of Health and Human Services (DHHS) issued regulations interpreting the community service and uncompensated care requirements. These regulations placed severe time limits on the uncompensated care provision, but did not place any time limits on the community service requirement. For most Hill-Burton hospitals, the uncompensated care requirement expired more than ten years ago, having had only a limited effect on the quantity of uncompensated services provided to indigent patients. The courts have consistently rejected the opportunity to expand Hill-Burton’s community service requirement, which remains in place, beyond the DHHS regulatory provisions.

The community service obligation of the Hill-Burton Act is one of the clearest opportunities for courts to interpret legislation to reduce inequalities in access to health care services. Even though the regulation permits denial of care based on ability to pay, there is wide latitude in the regulation for interpretations that could dramatically increase the availability of health care services to indigents. But the judiciary has declined this opportunity. Courts have not taken advantage of the community service provisions to require Hill-Burton facilities to address inequalities. There are few (if any) reported cases that have used the community service requirement to reduce disparities in health care services. Tellingly, in a 2000 volume devoted entirely to health inequality in the U.S., Hill-Burton is not even mentioned as a potential legal remedy.

**ERISA**

The Employee Retirement Income Security Act (ERISA) offers an interesting study in the doctrine of unintended consequences. Originally enacted to pre-
In strongly deferring to Congress, the Supreme Court has sent an unmistakable signal that it does not view its mandate as alleviating market deficiencies or inequalities.

The Culture of Technology

Another item on my post-tort reform agenda is to address the effects of the technological imperative on health care delivery/policy and legal doctrine. My working hypothesis is that technology drives medical liability and health policy. Although data are not available to determine whether a few technologies account for a significant portion of medical liability risk, or to quantify technology’s overall contribution to claim frequency and award severity, it seems clear that technology is the major driver of medical liability trends. But more significantly, the nation’s culture of technology underlies the relationship between law and medicine. No other factor plays such a powerful explanatory role in litigation trends or overall health policy. Observers of the U.S. health care system frequently remark on the nation’s culture of technology. Americans expect, indeed demand, both continued innovation and widespread (though not universal) availability. For reforms of both the health care delivery system and the medical liability system to be effective, we must contend with the culture of technology.

Historian Kenneth De Ville has been a leading proponent of the relationship between technology and medical liability. His explanatory framework for recurrent medical malpractice crises invokes both long-term cultural factors and short-term topical influences, and is a useful starting point for understanding litigation trends and for adopting appropriate policy responses.

Under long-term cultural trends, De Ville notes several factors: 1) an upward-sloping baseline proclivity to sue; 2) breakdown of community solidarity that discouraged litigation; 3) a rising secular belief that humans can improve their lives; 4) a growing preoccupation with physical well-being; and 5) increased demand that there be a remedy for every wrong. Topical influences include: 1) attitudes toward the medical profession; 2) more sophisticated plaintiffs’ attorneys; 3) increasing media coverage; 4) changes in legal doctrine; and 5) the absence of national health insurance. The inadequacy of health insurance deserves special attention as an incentive for litigation. When technology-laden medical care is needed following iatrogenic injury, the expensive uninsured patient can be devastating.

Three aspects of the cultural dimension particularly influence the law-medicine interaction. One is the oft-noted phenomenon of unrealistic expectations — that all medical interventions, particularly those relying on innovative technology, will be successful. A second aspect is the pressure that cultural expectations put on manufacturers and physicians to use the latest technology without adequately assessing its value. The ex-
amples of electronic fetal monitors (EFMs) and high-dose chemotherapy with autologous bone marrow transplant (HDC-ABMT) for metastatic breast cancer patients show the dangers of premature technology diffusion. The third aspect is overconfidence in the scientific basis of technological innovation, which reinforces the lack of assessment by not putting pressure on the system to justify new technologies.

Any enduring solution to the recurrent medical malpractice crises and to the costs of health care must include sustained engagement with the underlying culture of technology. I suggest that we need a much more forthright debate about the cultural aspects of technology, with physicians and attorneys, along with academics, taking the lead. It may well be that we have simply become too devoted as a society to the technological imperative to have a meaningful dialogue about it right now. In fact, there are powerful arguments in its favor. There is no question that technology permits physicians to attack conditions that would otherwise cause suffering and even death. Technology has clearly contributed to higher quality of life and perhaps longevity as well.55

Yet without some way to limit public expectations, physicians will never escape the quandary technology imposes. Too much pressure exists throughout the medical system to adopt innovative technologies without adequate time to determine their appropriateness and effectiveness. The current fragmentation in the health care system, with profitable specialty facilities advertising the latest in technological advances, forces competitors to match the technology. So a technology spiral is evident: patients want the latest technology (especially when they do not absorb the full costs), physicians are forced to provide it or else lose patients, and manufacturers are quite content to oblige with new products.

At best, cultural change will not occur quickly and will be difficult to achieve. To begin the process of altering the culture of technology, it seems to me that physicians can play a significant role by communicating the limits of technological solutions to medical problems. This needs to occur at both the individual patient encounter and at the professional society level. During the physician-patient encounter, patients need to understand that technology is not a magical solution. But patients whose only hope lies with the latest technology, even if unproven, have little incentive to listen to the caveats. At the professional society level, medical leaders should begin a dialogue with the public based on more realistic expectations about medical care’s limits. The dialogue should focus on reducing the public’s reliance on technology and properly depicting technology as only one aspect of medical care.

As a caveat to this discussion, there are serious issues of ethics and values that must be examined when placing limits on technology. When dealing with “last hope” interventions, it is understandable that individuals will want not just aggressive therapy, but the latest technology available (even if it has not been shown to be effective in clinical trials). In the dialogue about limits to technology, we must at least keep individual patient needs and legitimate demands in mind. Perhaps the best we can do is to manage conflicting values and develop institutions to mediate the competing interests.56

In sum, the technological imperative is an issue that should be part of the broader health policy and health law debate. In particular, health law scholars can contribute to a critical analysis of the conflicting values and how the law can best balance the competing interests.

**Mutual Distrust Between Attorneys and Physicians**

A third area that health law scholars should address is the toxic relations between attorneys and physicians.57 For reasons that are both historical and related to competing policy objectives regarding medical malpractice litigation, relations between the two professions have deteriorated substantially over the past two decades. If the effects on patient care and health care delivery were not so serious, we could laugh this off as more reminiscent of the periodic (if enduring) feuding between the Hatfields and McCoys than a serious long-term breach between two esteemed professions. Come to think of it, the periodic malpractice crises are beginning to resemble the Hatfields and McCoys!

Yet the animosity is serious and threatens to spin out of control. With some physicians now threatening to refuse care to attorneys and their families, the possibility of a permanent breach is no longer theoretical. But state level tort reforms may permit a dialogue over how to return to a more productive relationship. If so, health law scholars (especially those with ties to the medical community) can play a crucial role in the reconciliation process. In my view, the patient safety movement will not reach its full potential without some concordance between physicians and attorneys that allow for collaborative policy efforts centered on patient care, not on protecting professional interests.

**HEALTH LAW 2005 – TRENDS AND PROSPECTS: OR, WHY I’M ALWAYS WRONG!**

**Competing Paradigms**

In co-authoring a law school casebook with Larry Gostin, I argued that:

Four conceptual paradigms can be applied to transform the health care system: economic (the stan-
skeptical that markets will protect access to health care. My thinking about the competing paradigms was largely rooted in an ongoing struggle between market proponents (a consumer-driven health care system), proponents of a social justice model (largely governmentally determined), and medical professionals. In this framing, tort versus contract is a convenient proxy for the translation of the competing paradigms into legal rules.

Realistically, the social justice model is on hold—that’s a euphemism for being dead in the water. Instead, the real struggle for doctrinal supremacy in health law is between the market and professional models. The oft-studied role of tort versus contract in setting legal doctrine in health care delivery is simply one aspect of that larger battle. By both temperament and philosophy, I remain rooted in the social justice model, skeptical that markets will protect access to health care. Just as importantly, my view is that physicians remain central to the health care enterprise and are responsible, legally and ethically, for patient well-being. The reality is that clinical care starts and ends with physicians rather than with institutions. As a normative matter, the physician-patient relationship should be central to the health care enterprise. Unlike market-oriented scholars, therefore, I prefer a legal regime that is patient-centered—based on professional norms as opposed to market forces.

But preferences and realistic assessments of the direction of legal doctrine are not the same. Surprisingly, the professional paradigm seems more alive and well than others (including Professor Havighurst) would like (if less so than I would prefer). Take the battle between tort and contract for dominance in medical liability doctrine as an example. Without question, contract is an increasingly important force. Contractual arrangements requiring arbitration have generally withstood legal challenges. Yet for the most part, deference to professional norms remains the standard in determining medical liability. Even if one accepts Professor Philip Peters’ analysis that a trend among state courts is to move away from deference to the professional custom standard toward a “reasonable and prudent physician” standard, it is not clear how the emerging reasonable physician standard differs conceptually from professional custom and whether case outcomes are actually different in jurisdictions switching to the new approach.

Likewise, antitrust doctrine, at least with regard to quality of care considerations, has not fully adopted a market model. As noted above, recent cases have opened the possibility that antitrust courts will rely on professionals to define quality of care. Whether doing so would limit the effectiveness of antitrust enforcement remains to be seen.

Governance may be a third area where professional norms will play a more pronounced role. Health care administrators’ fiduciary duties are so suffused with medical professional norms that courts may well defer to them as opposed to business norms. True, executives will be accorded great latitude in exercising their best business judgment. Even so, it seems difficult to extricate business judgment from medical professional judgment, especially in a nonprofit institution. If the Greaney/Boozang mission primacy standard holds, it seems inextricable from medical professional norms as a way of defining what constitutes mission priority.

Future Prospects
One critique of the foregoing analysis is that it selectively highlights a few instances of professional norms...
being upheld and implicitly extrapolates that into a sustainable doctrinal trend. That’s probably an accurate assessment. Indeed, it would be foolish to suggest that courts are likely to protect physicians at the expense of surging market forces. Most likely, the inexorable trend is toward legal doctrine in health care that supports markets. In fact, I think a reasonably compelling case can be made that legal doctrine has already developed to support the market arrangements through which managed care has become dominant. With the Supreme Court’s recent decision in the Davila case,60 returning to the Court’s original broad ERISA preemption doctrine, I see little likelihood that legal challenges to managed care’s preeminence will be successful any time soon. At least in health care, the courts seem to be willing to let the market determine the winners and losers.62 Nevertheless, even the examples of professional norms cited above suggest that courts may be reluctant to abandon physicians altogether to market arrangements.63 Why that is so I’m not entirely sure. One possibility is that judges have the same reflexive preference for physician-centered care (as opposed to institutional-centered care) as I do. Another is that the academic dominance of law and economics, which would almost surely result in limiting professional dominance, has not yet taken hold in judicial decisions. Professor Gregg Bloche expands on the role of professional ethics and norms, arguing that they pervade health law. He observes that while antitrust law may be the most visible conflict between professional norms and market arrangements, this tension is manifest in many issues. “Conflicts over the lawfulness of financial rewards to physicians for frugal practice, the authority of treating physicians versus health plan managers to determine medical need, and the supervisory powers of plan managers over clinical practitioners pit professional norms against immediate market pressures.”64 This analysis suggests that it may be difficult for courts to ignore the role of professional norms, even if judicial doctrine favors market arrangements.

To be effective, professional norms need to adapt to changing marketplace realities without sacrificing the core value of patient advocacy. As Professor Gail Agrawal argues, professional norms are essential for legitimizing cost control programs through cost-conscious clinical decisions. “Professional norms are also part of a social contract between the medical profession and the public,”65 and can comfortably exist within a framework dominated by the economic model.

**CONCLUSION**

Health care is a field driven by fads. Just a few years ago, managed competition was the solution to the system’s deficiencies. Then it was health insurance purchasing cooperatives, followed by business health purchasing coalitions. Along the way, managed care emerged as the ultimate solution. Each of those was exposed as flawed. Now the mantra is consumer-driven health care. That, too, will be exposed as flawed and another fad will emerge. All of this makes it difficult to establish sustainable legal doctrine in health care. Health law scholarship might profitably focus on some of the areas outlined above that will be critical for the relationship between law and medicine regardless of what model of health care delivery emerges over the next few years.

Even if health law does not lend itself to the development of legal theory that drives the legal academy, the more prosaic aspects of developing sound legal doctrine in health care are critically important for practitioners. As an applied field, health law scholars have much to offer judges, policymakers, and, most importantly, health care practitioners.

---

**References**

10. Hall and Schneider, supra note 8, at 103.
11. I am not a supporter of caps on damages, the primary reform the AMA desires, and doubt that caps will make much of a difference to malpractice insurance premium rates.
dard would be whether there is a change in corporate purpose that requires state attorney general or judicial approval, usually through a cy pres proceeding, a far more stringent test than the good faith and reasonable business judgment standard of fiduciary duty. Critics of applying charitable trust principles argue that it imposes too high a standard and would impede trustees’ ability to respond to market changes. To date, the courts have not established either one as the dominant standard. See M. W. Peregrine and J. R. Schwartz, “Nonprofit Health System: New Legal Developments Suggest Alternative Corporate Structures,” Health Law Digest Analyses (2004) (criticizing the use of charitable trust principles to challenge nonprofit trustee decisionmaking authority, hence in agreement with Professors Greaney and Boozang).


18. See, e.g., M. A. Hall and J. D. Columbo, “The Charitable Status of Nonprofit Hospitals: Toward a Donative Theory of Tax Exemption,” Washington Law Review 66 (1991): 307 (arguing that providing a community benefit does not justify tax exemption, and that there is no causal connection between the exemption and the community benefit provided); see also M. G. Bloche, “Health Policy Below the Waterline: Medical Care and the Charitable Exemption,” Minnesota Law Review 90 (1995): 299-352, at 350 (arguing that tax exemption can be retained, but proponents have failed to justify it on functional grounds because of three primary failings: “insufficient provision of charity care to the uninsured poor; reluctance to treat Medicaid patients, and inadequate provision of health promotion, clinical screening...and other outreach services to needy communities.”).

19. But see Horwitz, supra note 17 (arguing that nonprofits provide a disproportionate array of public benefits beyond charity care than for-profits).

20. See Bethesda Healthcare Inc. v. Wilkins, 806 N.E.2d 142 (Ohio 2004) (holding that a fitness center that the plaintiff nonprofit hospital owned was not entitled to a tax exemption because few non-hospital members used the center); see also NIC Health Plans, Inc. v. Commissioner of Internal Revenue, 325 F.3d 1188 (10th Cir. 2003) (upholding the county's refusal to grant tax exempt status because the system failed to meet the community benefit standard).

21. Nevertheless, Professor Jill Horwitz has found that nonprofits still retain more unprofitable services than for-profits. See Horwitz, supra note 17.


27. 65 F.3d 1406 (7th Cir. 1995), cert. denied, 516 U.S. 1184 (1996).


33. See, e.g., W. M. Sage, “Understanding the First Malpractice Crisis of the 21st Century,” in A. M. Gofshel, ed., Handbook of Health Law (2004) (noting the historical irony that physicians argued against medical liability because errors were few while now supporting efforts to keep errors confidential).

34. See, e.g., Institute of Medicine, To Err is Human: Building a Safer Health System (Washington, DC: National Academies Press, 2000); Institute of Medicine, Crossing the Quality Chasm: A New Health System for the 21st Century (Washington, DC: National Academy Press, 2001).


41. Rosen v. Goetz, 2005 U.S. App. LEXIS 9743 (6th Cir.); see also Slalala v. Grijalva, 526 U.S. 1096 (1999) (holding that a Medicare carrier was a private entity and not subject to due process requirements). Cf. Harris v. Board of Supervisors of Los Angeles County, 366 F.3d 734 (9th Cir. 2004) (holding that the county could not close a rehabilitation facility because it would result in irreparable harm to indigent community residents).


48. In fairness to the judiciary, the absence of federal regulatory enforcement of the community service provision is a substantial reason why courts have not been more involved.


52. The subsequent case of Aetna Health Inc. v. Davila, 124 S. Ct. 2488 (2004), basically preempting most state tort litigation under ERISA, only compounds the problem.


55. See, e.g., D. M. Cutler, “The Potential for Cost Savings in Medicare’s Future,” Health Affairs 2005; Web Exclusive, W5-R77 -http://content.healthaffairs.org/cgi/content/abstract/hlthaff.w5.r77-


57. For an expansion of this argument, see P. D. Jacobson and M. G. Bloche, Improving Relations Between Attorneys and Physicians, Journal of the American Medical Association (forthcoming 2005).


59. A clear exception to this is that battles over tax exemption doctrine remain disputes between market and social justice models. Under a market model, only the private sector can “save” charity hospitals by converting them to for-profit status.


62. For a more detailed analysis, see Jacobson, supra note 49.


65. Agrawal, supra, at 381.

Introduction

Pain Management in the Emergency Department: Current Landscape and Agenda for Research

Sandra H. Johnson

This symposium issue on Pain Management in the Emergency Department: Current Landscape and Agenda for Research is the fifth special issue on improving care of patients in pain published by the *Journal of Law, Medicine & Ethics* with the support of the Mayday Fund. Since the beginning of the series in 1996, these *JLME* issues have been the source of original research on issues that impede access to care for patients in pain. These volumes have also had a significant impact on the course of public policy. As detailed in the foreword to the 2003 volume, the research done by the Mayday Pain Relief Project at the American Society of Law, Medicine & Ethics, funded by the Mayday Fund, has been key in the development of model policies and legislation to meet the needs of individuals in pain and the health care professionals who care for them.

With this volume, the journal series for the first time undertakes the task of mapping the landscape of treatment for pain in one particular health care setting. The emergency department is a perfect setting for this type of examination.

All stressors in health care converge in the emergency department, and some believe that the clearest window on the operation of the U.S. health care system is provided by the ED. This certainly appears to be the case in regard to treatment of patients in pain.

Emergency medicine carries the primary responsibility for treating patients who are suffering from traumatic injuries, but our system also relies on the emergency department for treatment of chronic pain, as well as pain at the end of life. All of the issues that affect the quality of care for suffering patients in the majority of health care settings are on the table in the ED: concerns about reliance on mistaken customary standards in treating (or neglecting) patients in pain; limitations on resources; discontinuity of care in complex and often chronic conditions; chemical dependency; race-based differences in treatment; wariness of the potential for deceitful individuals to pose as patients; and a lack of hard data to identify the most effective interventions to change physician behaviors or even the most effective treatments for pain in particular circumstances, among many other issues. Of course, each health care setting faces some unique circumstances, but one reading this symposium issue will understand that a great number shared critical factors present themselves in the ED.

Even if the emergency department were not the microcosm that it appears to be, the problem of inadequate treatment for pain in the ED has significant consequences in and of itself. Studies referenced in the articles that follow indicate that isolated episodes of acute pain can increase pain in later acute events, can impair recovery from trauma, and can trigger the development of a chronic pain syndrome.

We hope that this symposium stimulates further research on the treatment of pain in emergency medicine. Everyone has a stake in improving care in the ED. Those who advocate for chronic pain patients know that the ED is a critical part of their care, either because unavoidable exacerbations may occur or because treatment for those patients in the office-based setting may be less than optimal. Those who are concerned about improving care for the dying understand that many patients with terminal illnesses will end up in the ED for acute events or may be taken to the ED when death is imminent. Even those focused on improving care in nursing homes are aware that the emergency department has a significant role to

*Sandra H. Johnson, J.D., LL.M., holds the Tenet Chair in Health Law and Ethics at the Center for Health Law Studies and the Center for Health Care Ethics at Saint Louis University and directs the Mayday Project at ASLME.*

PAIN MANAGEMENT IN THE EMERGENCY DEPARTMENT • WINTER 2005

739
play, especially when transfer to hospitals is so common for nursing home residents who are dying. Frankly, everyone knows that their child or their spouse or they themselves may one day be dependent on an emergency department in a time of pain and suffering.

As always, efforts like this volume rely on the generosity and commitment of a handful of people. First, we are deeply grateful to the Mayday Fund who supported the initiative on pain relief in the emergency department. This volume represents only a small part of their commitment to this effort. In October, 2004, for example, the Mayday Fund provided support for a national leadership meeting in San Francisco sponsored by the ASLME in conjunction with the annual meeting of the American College of Emergency Physicians. That meeting – “Toward the Painless ED: Better for our Patients, Better for Us” – convened leaders in emergency medicine, patients’ advocates, and scholars and public policy experts in related fields to learn about the challenges of pain relief in the ED and to identify priorities for research and action. That conference informed much of the subsequent research that you are reading in this volume today. In addition, in line with its commitment to make sure that research on pain makes a difference to patients in pain, the Mayday Fund is supporting the production and distribution of this volume to opinion leaders who can influence emergency medicine practice.

The Mayday Fund’s consistent, and persistent, support of the pain relief projects at the American Society of Law, Medicine and Ethics has been remarkable. We are grateful to the trustees of the Fund and to the Executive Director Christina Spellman for their passion and insight as well as their financial support. They actively monitor their grants to make sure that they are productive, and we have appreciated that collaboration.

This project also benefited from the advice and counsel of a knowledgeable advisory board. Members of the board included Jim DuCharme, Scott Fishman, Barry Furrow, Keith Ghezzi, Rob Schwartz, and Knox Todd. In particular, the guidance of Knox Todd and Scott Fishman, as we explored the world of emergency medicine, was absolutely essential and always provided with great graciousness and generosity. Knox and Scott are completing other Mayday-funded research on emergency medicine, and we are looking forward to publication of their results soon. We also acknowledge the significant assistance from the American College of Emergency Physicians. ACEP’s support for the San Francisco meeting was critical to its success, and its continuing interest in pain management will be essential to improving care.

Finally, the American Society of Law, Medicine & Ethics has been fully committed to the Mayday Projects for the past nearly ten years. Always going above and beyond original expectations in every case, ASLME, and especially Ben Moulton, the Executive Director, devoted care and resources to making sure that the projects exceeded their goals. The Mayday Project was never just another event, or just another volume of the Journal.

Acknowledgements
The author gratefully acknowledges the support of the Mayday Fund for this research as well as the research assistance of Carine M. Boyle.
The Social, Professional, and Legal Framework for the Problem of Pain Management in Emergency Medicine

Sandra H. Johnson

Context and Challenges in the Emergency Department

The problem of harmful, unnecessary and neglected pain has been studied extensively in many health care settings over the past decade. Research has documented the incidence of untreated pain, and scholars and advocates have given the problem several names: “public health crisis,” “oligoanalgesia,” and “moral failing,” among them. Articles have identified a litany of now familiar “obstacles” or “barriers” to effective pain relief. Each of these individual obstacles or barriers has been the subject of targeted remedial action in at least some context.

The checklist approach to improving care for patients in pain, however, is likely to have only limited effect. What really appears to be operating is a complex ecosystem that supports ambivalence, denial, and even suspicion of the circumstance of patients in pain and efforts to treat them. Pain relief in emergency medicine, a relatively new setting for the study of challenges to treating pain, provides a revealing context for viewing discrete obstacles to effective pain management in medicine as part of an integrated environment into which patients with pain enter for treatment.

It is pain that drives most patients to seek care in an emergency department. In the majority of patients, the pain that drives them is quite severe, rating an 8 of 10 on commonly used pain scales. Emergency medicine, however, does not focus on the management and relief of pain. Pain is most commonly, and necessarily, viewed as a symptom that guides the physician to a diagnosis of an underlying pathology. It is only when pain is viewed merely as a symptom, rather than a pathology itself, that there is a problem.

The model of pain as merely a symptom does not serve a good number of patients coming to the ED with pain. In fact, a significant proportion of emergency patients suffer serious and debilitating chronic pain; and approximately 11% of patients seeking treatment in the ED do so for pain related to chronic conditions. For persons with chronic diseases associated with acute episodes of pain, including sickle cell and migraines for example, the sole purpose for the visit is pain relief. Although active diagnostic efforts may still be necessary to rule out other conditions that may be causing this particular pain episode, the treatment of the pain itself is obviously the primary objective of emergency treatment in those cases.

Sandra H. Johnson, J.D., LL.M., holds the Tenet Chair in Health Law and Ethics at the Center for Health Law Studies and the Center for Health Care Ethics at Saint Louis University and directs the Mayday Project at ASLME.
Even with the need to focus on diagnosis and treatment of the condition causing the pain, however, pain management and pain relief should be a priority in emergency medicine. The ethical duty to relieve pain is well established. Although there may be ethical and medical concerns about particular pain management interventions in particular circumstances, the core ethical obligation to relieve pain is well established in medicine, including in emergency medicine. The Code of Ethics of the American College of Emergency Physicians, for example, formally recognizes an obligation to relieve pain as a part of emergency treatment. Relieving pain and suffering has been called a “fundamental imperative for any clinician”; and, in regard to emergency medicine: “as a guiding principle of medicine and core covenant with our patients, every EP [emergency physician] must embrace providing timely and effective pain control as a fundamental duty.”

Other emergency physicians have observed that “[o]pportunities to save lives within the ED are rare, but opportunities to relieve pain are nearly infinite....” These statements represent specific application of more general ethical norms to the particular context of emergency medicine.

In addition to ethical norms that support serious attention to pain management generally, there are pragmatic reasons, specific to emergency medicine, that support an emphasis on pain relief as a priority. One might mistakenly view pain associated with an emergency condition as a temporary, though serious and intense, experience. Studies on the relationship between chronic pain and acute pain episodes, however, indicate that an experience of unrelieved, acute pain can make a person vulnerable to a pattern of chronic pain or to a repeat pain episode. Studies have also indicated that managing pain post-surgically promotes recovery, while persons with untreated pain are more likely to experience complications after medical treatment. It seems reasonable to extend these findings to untreated pain caused by trauma or non-surgical but painful medical procedures as well. Finally, one might speculate that a patient’s experience with painful procedures could lead that person to delay or avoid necessary medical diagnosis and treatment of a later episode or a new symptom. In fact, there is some evidence that a procedure may be more painful the next time it is employed if adequate analgesia was not provided in the first instance.

Despite the clarity of the ethical principles and the documented outcomes of untreated pain, research on the treatment of pain in emergency medicine has revealed a pattern of inadequacy. In 1989, Wilson and Pendleton applied the term “oligoanalgesia” to the neglect of pain in the ED and documented that 56% of patients in the ED presenting with pain received no analgesia; furthermore, when narcotic analgesics were provided, they were provided in doses too low to be effective. In particular, there is evidence of disparities in the treatment of patients for pain in the emergency department based on race and ethnicity. This evidence mirrors racial disparities in the assessment and treatment of pain in medicine generally. Similarly, studies have demonstrated that children receiving treatment in the ED are much less likely to receive pain medication for clearly painful conditions as compared to adults presenting with the same conditions. As in the case of racial disparities, the evidence of neglect of treatment for pain in children in the ED parallels identified problems in the care of children in other health care settings. Because approximately one-third of ED visits involve treatment of children, this has been a significant concern, addressed aggressively in some hospitals.

Empirical research on the reasons for the neglect of pain in the emergency department is quite thin. The pace of publication on issues of pain management in the ED increased significantly between 1996 and 2003, however, and appears to have continued an exponential growth since that time, indicating a promise of more research to come.

Despite the increase in attention, the problem of undertreatment of pain in the ED persists. As in other areas of medical practice, institutional initiatives in emergency medicine, including educational interventions and the establishment of departmental protocols to improve the treatment of pain in EDs, although sometimes successful, have very often produced disappointing results. Similarly, clinical guidelines on pain management standing alone have not been proven effective in changing physician practices. The lack of strong success in these efforts may be attributed to the design or implementation of the specific intervention. For example, Ducharme, in his article in this symposium, notes that practice guidelines are more
effective when there is personalized follow-up and mentoring with physicians in their own practice. The pattern of disappointing results in some of these remedies to identified barriers (for example, clinical guidelines responding to deficiencies in knowledge base), however, may reveal instead that the reasons and root causes of undertreatment of pain in the ED are still not well understood.

There is a more substantial literature on barriers to effective pain management in other areas of medical practice. The obstacles identified generally in medical practice include financial restrictions, educational deficiencies, cultural challenges, and legal and regulatory concerns, among others. It is likely that some of the reasons for undertreatment of pain in the ED are the same as those for medical practice generally. For example, some observers and practitioners have identified deficiencies in the educational programs that prepare emergency physicians.

Further, while financial issues, including payment and reimbursement for care, have been identified as significant barriers for pain patients outside of the hospital setting, the emergency department faces different financial issues. For emergency medicine, financial constraints are often expressed in terms of capacity relating to patient load and crowding. One might expect that the volume of demands on the ED negatively impact attention to pain management. At least one study, however, has indicated that staff-patient ratio (weighted by acuity of the patients’ conditions) did not affect the proportion of patients who received pain medication.

As in other settings, institutional structure and procedures may also form barriers to effective pain relief in the ED. For example, ED procedures, typically requiring at least seven steps (“patient presentation and registration, nursing assessment and triage, placement in a treatment room, primary nurse assessment and documentation, physician evaluation, physician ordering of pain medication, nursing obtaining pain medication, and finally, nursing administration of pain medication”) before the patient can receive any pain medication, create a formidable barrier to timely treatment and the avoidance of unnecessary suffering. Several studies have documented lengthy delays in the first administration of pain medication to ED patients suffering serious trauma, and studies of patient’s expectations indicate that these delays are probably a source of significant concern to patients. Pre-hospital emergency medical services have also been identified as a target for improvement of pain relief for the emergent patient.

The practice of emergency medicine is quite different from other areas of practice, however, and some of the reasons for neglect of pain may also be distinctive. Because there is little empirical research on obstacles to effective pain management in the ED, most of the reasons given for the phenomenon in emergency medicine emerge from intuition and experience or are extrapolated from the few studies that exist. Further research is certainly required, but some preliminary conclusions are possible.

Distinctive reasons for undertreatment of pain in emergency departments include the prioritization of diagnosis over pain relief; inadequacies in the process of pain assessment; and a culture that supports significant detachment from patients. Recent literature has identified legal risks as an additional cause of concern for emergency physicians. Areas of liability risk, including litigation over recklessness in the neglect of treatment for pain, spotlight systemic issues that impact the quality of treatment for pain in the ED. These include discontinuity of care, especially relating to arranging for the treatment of pain upon discharge as well as inadequate pain management by providers outside the ED; challenges of palliative care in the ED; limitations on the scope of practice of emergency health care professionals that affect the timeliness of pain management interventions; and issues around informed consent. In addition, no discussion of emergency departments would be complete with consideration of the application of the federal Emergency Medical Treatment and Labor Act (EMTALA) to the question of treatment for pain. Finally, even though emergency physicians work in a different legal environment than does the doctor in an office-based practice, they may share some concern over the risk of regulatory action for the prescription of controlled substances.

**Subordination of Pain Relief to Diagnosis**

The subordination of pain relief to diagnosis in emergency medicine is likely to be one reason that interventions to relieve pain are delayed or denied in the emergency department. The emergency physician’s priority is diagnosis. Patients share this priority for diagnosis and treatment, and some evidence indicates that patients thus may simply expect to experience suffering in the emergency department resulting in lower patient demands for analgesia. Other studies, however, indicate that patients have substantial expectations for pain relief in the ED.

Treatment that addresses only the symptoms of pain and neglects the underlying cause is recognized as substandard emergency care. Because both the patients and the physicians in the context of an emergency desire accurate diagnosis above all, and because emergency medicine is held to a medicolegal standard that holds them accountable for negligence in diagnosis, it
is important to address the pain-diagnosis connection as an empirical question. There has been a strong belief in emergency medicine that analgesia may impede diagnosis, and that belief has impeded pain management in the ED. Where evidence can be produced to reject the hypothesis that interventions to relieve pain hamper diagnosis, the practice of withholding analgesia in favor of diagnosis should be expected to change. In such situations, one would not need to argue whether pain relief was worth a reduction in diagnostic efficacy because the two concerns would not, in fact, conflict.

The case of pain management in the context of a patient presenting acute abdominal pain (often called “an acute abdomen”) is illustrative. The well established practice and custom in emergency medicine has been to withhold pain medication from persons with acute abdominal pain until diagnosis or surgery. This practice reflects the long-held belief that interventions to relieve acute abdominal pain would confound the diagnosis of the underlying cause.

The firmly held, persistent “common knowledge” in emergency medicine that medication of acute abdominal pain would thwart accurate diagnosis is based on a statement in a medical textbook from the 1920s. In that text, Sir Zachary Cope taught: “though it may appear cruel, it is really kind to withhold morphine until one is certain or not that surgical interference is necessary, i.e., until a reasonable diagnosis has been made.” Similar statements were included in this textbook as recently as 1979. Even though Sir Zachary provided no evidence or citations for his statement, “physicians everywhere took Dr. Cope’s opinion on this as their own.” Sir Zachary’s opinion has guided doctors for more than 80 years, one generation after another, despite the fact that no study supported the practice. A study published in 1998, for example, revealed that emergency physicians in the U.S. withheld analgesia for patients with acute abdominal pain even though Sir Zachary’s conclusion had been challenged in the literature nearly 20 years before. A series of studies, each concluding that treating the patient with acute abdominal pain with morphine did not impede and perhaps even enhanced the accuracy of diagnosis, were published in the 1990s. Studies published since that time also failed to detect any adverse impact on diagnosis.


Administration of narcotics to patients with abdominal pain to facilitate the diagnostic evaluation is safe, humane, and in some cases, improves diagnostic accuracy. Incremental doses of an intravenous narcotic agent can eliminate pain but not palpation tenderness. Analgesics decrease patient anxiety and cause relaxation of their abdominal muscles, thus potentially improving the information obtained from the physical examination. There is evidence that pain treatment does not obscure abdominal findings, or cause increased morbidity or mortality.

In the face of Sir Zachary’s warning, unsupported by any evidence, and of early studies that failed to document an adverse impact on diagnosis and instead produced evidence contrary to this traditional “knowledge,” the ACEP panel responsible for the Policy chose to rely on early evidence, such as it was, rather than perpetuate the traditional practice of leaving the patient to suffer, a practice that was supported in none of the studies available at that time or since. ACEP’s Policy recommendation regarding pain treatment for this category of patients is clear and firmly stated. The only evidence that exists does not support the customary withholding of narcotics.

This portion of the Policy, however, is categorized as an “option” rather than an “evidence-based standard” or “guideline.” Recommendations in the Policy based on a “high degree of clinical certainty” and supported by the highest level of empirical research qualify as an “evidence-based standard” in ACEP’s policy. In contrast, a recommendation labeled an “option” is a “strategy” for patient management based on preliminary, inconclusive, or conflicting evidence, or, in the absence of any published literature, based on panel consensus. The Policy cites four research articles published prior to 2000 to support the option recommending attention to pain relief for ED patients with acute abdominal pain.

This level of uncertainty is not unique in medicine generally or in emergency medicine in particular. In the same Policy, for example, ACEP could not identify evidence-based standards or even guidelines for the diagnosis of several disease states that produce “commonly missed diagnoses” or for the diagnosis and management of geriatric patients as a subset of high-risk patients. Medical research of any sort is difficult in emergency medicine, and federal regulations have been developed to set up a process that accommodates some of those difficulties. Research on pain management in this situation in the emergency department is doubly difficult, both ethically and legally, because the construction of a control group would require some number of patients to suffer without pain medication in a...
situation where no current studies give any indication that medicating for pain relief has an adverse effect. Even with its limitations, one would expect the ACEP policy and the published studies to change practice. Some evidence, however, indicates that ED physicians still withhold analgesia for acute abdominal pain, illustrating the difficulties in changing embedded professional custom, even when no evidence supports the practice, as well as the sometimes slow diffusion of clinical knowledge. Another study, however, indicated that the Policy, or similar efforts, may have had an effect, although the study is ambiguous in an important respect. In a 2002 survey of emergency departments, 59 of 60 departments completing the survey responded affirmatively to a survey question that asked: “Is it your practice as a department to ever administer narcotic analgesia to acute abdomen patients prior to a surgeon’s evaluation?” A positive response to this inquiry does not indicate that it is common practice to do so, of course. The survey also asked: “in the cases when you do this, which of the following are motivations behind why you do it?” Of the reasons given, 88.1% reported that they would provide medication to alleviate patient discomfort; 86.4% believed that the literature supports the practice; and 61% responded that “it often takes too long for the surgical consult to arrive.” This latter question does indicate that there is a general awareness of the direction of current research on the matter. If clearer studies document that the practice of withholding analgesia in cases of acute abdominal pain does persist because of concerns over the impact on diagnosis, it may prove that we must look elsewhere for the reasons that pain relief is being withheld.

The implementation of ACEP’s policy on pain relief for this set of patients illustrates another distinctive factor in improving practices in the ED; i.e., the relationship of emergency department physicians to the other specialists on whom they must rely for treatment of their patients. As recently as 1996, 89% of surgeons surveyed in a single-state survey responded that they still preferred that such patients receive no medication for pain prior to the surgical consult. This survey of surgeons occurred less than two years before the survey in which 86.4% of emergency physicians indicated that the literature supported the use of pain medication in this situation. Further illustrating the conflict between these two specialties is a 2003 article in the American Journal of Surgery arguing that the studies that detected no difference in diagnostic effectiveness were infirm and that, based on cases “reported anecdotally and our own experiences,” analgesia can alter the physical examination and lead to misdiagnoses. The authors, therefore, recommend that analgesia should be administered “only with the knowledge and consent of the surgeon who assumes responsibility for decision-making.” This recommendation goes in the opposite direction of the ACEP Policy, and builds in significant delay in the treatment of patients with pain that may be harmful if acute pain is taken seriously and, at least under the current level of research, is unnecessary.

Pain Assessment, Or You Don’t Necessarily Know It When You See It

Recognizing pain, and understanding its severity, is not a simple question of empathy borne of shared experiences. In fact, experience of physically painful incidents or stimuli is not shared. There is great variability in the individual experience of pain in like circumstances. This variation is demonstrated in conflicting assessment of pain even by close intimates of the patient. Pain management in nursing homes, for example, is challenged by the tendency on the part of both health care providers and family members to underestimate pain in the elderly. Children also suffer from the inclination of adult caregivers, even family, to discount their reports of pain. Adequate pain assessment may require significant time, and assessing pain in individuals with cognitive impairment, a situation confronted with some frequency in the ED, requires even more effort. Further, strongly held assumptions that particular groups of patients, such as neonates, do not feel or remember pain have proven mistaken; and this seems to be the case with the current widely held assumption that sedated patients receiving emergency care do not feel or remember pain.

Perceptions of patient and physician as to the degree of pain experienced or expected are also often seriously divergent, including in the context of emergency treatment. Formal pain assessment techniques are intended to give voice to the patient in detecting the presence and severity of pain in a way that is informative to the health care professional and can lead to appropriate interventions to relieve pain. The importance of pain assessment is evident in the fact that the Joint Commission on Accreditation of Healthcare Organizations has incorporated required pain assessment as a linchpin in the efforts to improve care of hospital patients, including specifically care provided in the emergency department.

Many studies have identified deficiencies in pain assessment in the ED. Initial studies indicate that pain assessment is ordinarily a one-time evaluation in the ED and is not performed, or at least recorded, at important points after the initial assessment. In particular, assessment at the point of transfer or discharge, as discussed in the context of legal risks below, is critical to satisfying the emerging standard of care.

Pain assessment is known to be particularly difficult
where the patient is unable to communicate, as can often occur in the ED because of the patient’s age, mental disorder, trauma or stress. In addition, the emergency patient’s own emphasis on diagnosis and the resultant expectation of pain and suffering, noted above, probably makes pain assessment more difficult in the ED in the absence of formal inquiries of the patient by the nurse or physician and assessment techniques designed to elicit the patient’s perceptions. This assumption is based on a similar phenomenon described in the case of elderly patients who are observed to under-report pain for fear of being a burden. In addition, it appears that the expectation of pain can influence under-reporting of pain.

Intuitively, it seems that effective pain management must begin with recognizing the presence of pain. Indeed, an increase in the employment of analgesia for pain in the ED once a formal pain assessment system is adopted has been demonstrated. While it seems obvious that effective pain management in the emergency department requires formal pain assessment, especially in light of the divergence of perceptions between emergency physician and patient, there is some thought that the apparent connection between formal pain assessment and effective treatment is not so close.

Particularly troubling and challenging for the implementation of patient-directed pain assessment is evidence that emergency physicians interpret patients’ accounts of pain in a way that supports the physician’s assessment of the underlying situation. One study reporting highly variable responses among ED doctors to identical patient reports of a need for pain relief speculates that doctors who suspect that a patient is seeking drugs for other purposes will take a report of a need for pain medication as evidence confirming drug-seeking behavior, while physicians “who suspect the patient is truly in pain interpret the same statement [by the patient] as evidence that the patient is in severe pain.” In addition, physicians with more experience rather than less appear to be more likely to reject patients’ reports of pain, leading one author to argue that “[w]ithout ongoing education, senior physicians risk providing less, not more, pain control.

This phenomenon is not confined to the emergency physician or to the presentation of pain as physicians generally have been revealed to substitute their own assessment of the patient’s symptoms. Pre-existing interpretative frameworks for patients’ reporting of pain are particularly troublesome because of their influence on suffering and on public policy.

A Culture of Strangers under Stress?
Consideration of the culture of the ED, including the nature of the physician-patient relationship in emergency medicine, may reveal other reasons for under-treatment of pain. Emergency physicians and nurses work in a highly stressful environment where pain and suffering are immutable and relentless companions. They must act rapidly, with the understanding that their actions may jeopardize the patient’s life or health and in the face of intense uncertainty and unfamiliarity.

Influential research into the nature of the physician-patient relationship generally has discovered that “a central feature of doctor-patient interaction is the high degree of mutual uncertainty.” One has a sense that this may be exacerbated in the context of the emergency department where the doctors’ patients are strangers to them, and they to their patients. In addition, while a patient’s trust may bridge the inherent mutual uncertainty in the ordinary doctor-patient relationship, there is little basis on which to build trust in the ED encounter unless the patient has a reservoir of trust banked for hospitals and doctors generally. While some patients do come to the ED with that attitude, others come with the entirely contrary experience.

The emergency department typically serves patients who are strangers to the care team. The patient as stranger is so pronounced and profound that the issue is addressed specifically in the ACEP Code of Ethics for Emergency Physicians. The Code, in the section on Ethical Foundations of Emergency Medicine, specifically notes that “emergency physicians cannot rely on earned trust or on any prior knowledge of the patient’s condition, values or wishes regarding medical treatment.”

Although the statement specifically references the lack of knowledge about patient’s preferences, it also has obvious implications for pain assessment, most particularly for evaluation of honesty in the report of the patient’s pain.

There are other circumstances that also contribute to a more emphatic separation between physician and patient in emergency medicine as compared to other areas of practice. For example, emergency medicine is acutely aware of its role as providers of care to those persons whom everyone else has forgotten or avoids. This self-concept of rescue unit or safety net motivates professionals in emergency medicine to undertake the care of the abandoned and rejected as a part of their professional mission. It also speaks of a differentiation or even alienation from the patients served, however, and could contribute to difficulties in pain treatment.

There is evidence in other, non-emergency health care settings that patients with whom the physician is familiar receive more effective treatment for pain than do patients who are less well known to the doctor. It is reasonable to ask the question whether this phen-
nomenon is operational in the ED as well, because an answer to that question may produce significant insight into the problem of inadequate pain management.

In emergency medicine, the professional investment in favor of diagnosis itself also may produce an extreme form of detachment from the suffering of the patient who must be examined and treated. This detachment from patients in pain may in fact increase over time as the emergency department physician and nurse develop a tolerance for repeated and constant exposures to human suffering. The personality of individuals attracted to emergency medicine may personally discount the seriousness of pain and discomfort both for themselves and their patients. Maintaining a distance from the patient in pain may be a natural support for the need to proceed despite the patient’s suffering.

Unfamiliarity, detachment or alienation from patients may lead to a heightened fear of being tricked or duped by patients who have no medical need for controlled substances for pain relief. This challenge faces all physicians who treat a large number of patients in pain, but it is especially acute in the emergency department where the physician and the patient are usually unknown to one another. Experiences in which an individual takes on the mantle of “patient” but lies to the physician in order to get drugs seem to be nearly traumatic to emergency doctors and appear to breed a sense of betrayal and guardedness that can persist over the course of the physician’s career. Whether frequent or not, the experience is typically not an isolated incident for doctors in the emergency department. The problem is that the disgust at being tricked can become overgeneralized and result in the denial of necessary care to patients in pain.

When reasonable attention to this risk becomes fear, it leads to exaggerated distrust of patients’ reports of pain. A physician’s perception that a patient is seeking drugs for secondary gain is very powerful, so powerful that it may not be dislodged by anything the patient can do or say to alleviate that concern. As discussed above, at least one study has revealed that an emergency department doctor may, in fact, interpret an identical statement in polar opposite directions. The statement is interpreted as proving the doctor’s pre-existing perception whether that is proving that the patient is lying to get drugs or that the patient’s claim of pain and need is genuine. In addition, patients that fall within marginalized groups or groups that have been thought, based on evidence or not, to have higher incidence of diversion, may face a pattern of suspicion and limitations on care in a form of profiling.

Many efforts have focused on identifying indicators, often called “red flags,” that can be used in an attempt to cope with the possibility that some individuals may lie about their symptoms in order to get prescribed controlled substances. The usefulness of these efforts is questionable in a number of settings. The common “red flags” may be particularly unreliable when transferred from office-based medical practice to the hospital-based emergency medical practice. One commentator notes that ED doctors may be “ill advised” to rely on common “red flags” because these indicators have been developed in non-ED practice settings. For example, the request for a specific analgesic by a patient, commonly viewed as a red flag indicating a drug seeking patient, could indicate that the patient is suffering severe pain with which he or she is quite familiar. Another notes that the patient who has been discharged from the ED with pain medication (or a prescription for pain medication) who calls back or returns because the medication is “not doing the job” is suspected of abusing the system rather than suffering from inadequate dosing or selection of drug, problems that have been documented frequently in ED practice.

The ethical physician is alert to the patient who lies to get drugs for illicit purposes, but a serious ethical problem arises when the physician becomes hypervigilant or relies on profiling that gives only a general and often inaccurate picture of the “drug seeking” patient with the result that many patients in pain are denied necessary care. In fact, emergency physicians are likely to form suspicions about patients that are not influenced by the patient’s report of pain and that do not correlate with drug abuse screening.

When race, socioeconomic status, source of pay for care, and related generalities are used to exclude patients from effective treatment, ethical principles of medical practice are violated. The ACEP Code of Ethics is quite clear on the ethical principle involved. The Code states that “[e]mergency physicians should act fairly to all persons who rely on the ED for unscheduled episodic care.... Provision of emergency medical treatment should not be based on gender, age, race, socioeconomic status, or cultural background. No patient should ever be abused, demeaned, or given substandard care.” A situation where individuals are denied pain relief because of their health status (because they have sickle cell or because they are chemically dependent, for example) or because of stereotypes about a specific population implicates this ethical commitment.

The problem of physician distrust of patients is a core issue in the effective treatment of patients in pain. Increasingly, calls are made that doctors and nurses must “trust the patient’s report of pain.” Because of the high variability of the experience of pain and the impossibility in many cases of pinpointing an organic cause, the patient’s report is currently the primary, if not sole,
Legal Issues in Pain Management in the ED

Fear of legal risk has been identified as a significant barrier to effective treatment of patients in pain in a variety of settings. Concerns over the legal environment extend as well to the ED, although the source of these concerns is particular to this context. Legal issues relating to pain management in the emergency department emerge from at least three different areas of law. They are: 1) malpractice and general tort liability; 2) the federal Emergency Medical Treatment and Labor Act (EMTALA); and 3) state and federal regulation of medical practice, especially as it relates to the prescription of controlled substances. An analysis of legal issues relating to pain management in the emergency department is relevant because the sense of legal risk has an impact on the course of treatment. Perhaps more importantly in the context of the emergency department, an analysis of legal issues reveals systemic factors that may produce inadequate treatment for pain.

Malpractice and General Tort Liability

Physicians have a well-established legal duty to treat pain as a part of their medical treatment of a patient. The doctor’s legal duty to relieve pain is generally supported by policy statements and standards of professional organizations and by the standards enforced by state licensing boards. JCAHO standards on the assessment and treatment of pain in the emergency department also provide support for a legal duty to treat pain effectively. ACEP has adopted several policies that assert the importance of treating pain. A 2005 study reported that the National Guidelines Clearinghouse included 238 guidelines on “pain management,” including 143 guidelines on “acute pain management” as of December 2003. The courts rely on policy statements and practice guidelines promulgated by such organizations to establish a legal duty to which physicians and hospitals are held.

Litigation Concerning Negligent Treatment for Pain

Studies of malpractice lawsuits have concluded repeatedly that patients injured through negligence or malpractice generally do not file suit. In considering legal risks, efforts to improve pain management may be viewed pragmatically as a method for avoiding litigation, although this conclusion is largely intuitive. While undertreatment of pain is commonly viewed as an exacerbating factor in malpractice or negligence lawsuits, neglectful pain treatment standing alone can also form the basis of a malpractice or negligence claim. In Bergman v. Eden Medical Center and Tomlinson v. Bayberry Care Center, the surviving family members of two patients in California filed suit against the physicians, hospitals, and nursing homes that cared for the patients. In Bergman, the jury returned a verdict of $1.5 million, which the court reduced to $250,000. In Tomlinson, the defendants (the patient’s hospital physician, the nursing home physician, the hospital and the nursing home) entered into voluntary settlements with the plaintiffs, with undisclosed sums paid to the family.

Bergman and Tomlinson illustrate that it is possible for patients to bring suit for inadequate pain management in the absence of other negligence or malpractice. In each case, the patient was in the end stages of terminal cancer; the patient was transferred from hospital to nursing home for the final days or weeks of care; the patient received very clearly inadequate medication; and the lawsuits were both brought under the state’s elder abuse statute.

The diagnosis for each of these patients was clear, and the standard interventions for pain management were well accepted but were not provided. Treatment for cancer pain and pain at the end of life areas of treatment for pain in which there is a strong medical and legal consensus. There is no concern over addiction or diversion; and the state medical boards have long viewed the use of controlled substances for cancer pain, even over a long time and in large doses, as permissible. Lawsuits claiming neglected pain as the only basis for
legal action face several obstacles. Many states have a cap or limit on the amount of damages that can be awarded for pain and suffering. In some states, damages for pain and suffering do not survive the death of the patient and cannot be awarded to surviving family. It is for this latter reason that the plaintiffs in Bergman and Tomlinson brought their suits under a state elder abuse statute that provided a private right of action for elderly persons and their surviving family. Under this statute, however, the plaintiffs had to prove that the providers had been reckless and not merely negligent. This is a very difficult burden to meet in medical cases where professional judgment is so often the core of the issue. The statute provided for the payment of attorneys’ fees by the defendants to the plaintiff’s attorneys, and these fees amounted to approximately $500,000 in the Bergman case.

The threat of an avalanche of similar cases is not realistic because of the limits on this type of litigation. Furthermore, the facts of these cases as presented by the plaintiffs were quite extreme. Still, both Bergman and Tomlinson are particularly relevant to the practice of the emergency physician, even though at first glance they may be confined to terminally ill patients or patients with cancer pain. Their lesson is indeed broader, and highlights two common challenges to the quality of pain management for emergency medicine.

**The Risks of Discontinuity of Care**

The transfer from hospital to nursing home care in both Bergman and Tomlinson resulted in a serious discontinuity in care, especially at the point of discharge and transfer. This is evident in the absence of orders or follow-up for appropriate pain medication in at least one of the cases, despite documentation of the patient’s advanced cancer and consistent reports of extraordinarily severe pain.

Provisioning for adequate continuing pain management upon discharge from the ED is an issue for many types of ED patients. Several studies have identified serious concerns with failures to account for even basic pain management needs upon discharge.106 For example, a recent study of patients with orthopedic injuries, who were experiencing “acute distress” in the ED, revealed that 43 of 144 patients received no prescription or starter pack of medication upon discharge.107

Both Bergman and Tomlinson involved inadequate orders for pain medication upon discharge. The emergency physician must pay attention to transfer and discharge planning and assure that adequate medication and follow-up orders, including those required for pain management post-discharge, are provided for the patient.108 The ACEP policy on procedural sedation, for example, requires that continuing or developing pain and discomfort be addressed prior to discharge.109 Evidence suggests that EDs do not ordinarily document pain assessment subsequent to the initial assessment.110 An ongoing pain assessment in the ED is required for both treatment and discharge. Although there may be some question about the value of ongoing pain assessment during the course of treatment in the ED, it is difficult to understand how an appropriate post-discharge care plan for pain can be established without an assessment at discharge.

Another form of “discontinuity” of care presents a different kind of challenge to the physician practicing in the emergency department. Emergency physicians are familiar with the situation in which a patient who regularly receives care elsewhere for a chronic illness associated with pain comes to the emergency department for treatment for an exacerbation of their condition or for an acute pain episode. Emergency departments treat a significant number of chronic pain patients, accounting for more than one in ten of ED patients.111 The emergency physician is not as familiar with the patient as is the patient’s own physician, but it is the emergency physician’s services that are required.

An even more difficult situation occurs when the ED doctor is convinced that the patient is receiving inadequate treatment, for pain or otherwise, from their own physician or the facility in which they reside. In such cases, consultation with the patient’s doctor may help; serious and detailed information to the patient directly may allow the patient to take action; or admission to the hospital under the care of another attending physician may allow for more thorough assessment and a change in treatment plan.112

**Palliative Care for Terminally Ill Patients in the ED**

Bergman and Tomlinson both involved patients who were in the very end stages of terminal cancer. In particular, transfers of dying patients from nursing homes to the hospital, often through the ED, are frequent. At least one observer describes a “popular motto” in the nursing home world: “When in doubt, ship them out. Make the patient the other guy’s worry.”113 Emergency departments see cases in which a terminally ill patient who has been cared for in a nursing home or at home...
is brought to the emergency department when death is imminent. The admission of imminently dying patients through the ED presents challenges to the quality of care and pain management for these patients. Studies have indicated that the quality of care for such patients in the ED is not better than that received in the nursing home.114 Because the emergency doctors and nurses are not familiar with the patient’s medical condition or desires for treatment, interventions may be more acute than is desirable.115 Emergency physicians need to be familiar with the current practices and standards for effective treatment of pain at the end of life, and hospitals should have a plan to assure that the ED is well prepared to care for or admit these patients.116

Informed Consent and Pain Management in the ED
Informed consent is a serious challenge in the ED, and concerns over informed consent influence the effectiveness of interventions to increase responsiveness to patients in pain. This influence is seen in several areas: a concern that pain relief cannot be provided without informed consent; a “common knowledge” concern that opioids will disable patients from consenting to necessary interventions, especially surgery; and finally, a concern over potential liability for patients who take medications prescribed in the ED and who then engage in behaviors that are inadvisable because of the effect of the medications.

All medical care requires the informed consent of the patient, and medical treatment provided without consent is considered a battery. A limited exception to the requirement of informed consent exists in emergency situations. The classic statement of the exception for emergency treatment declares that the exception “comes into play when the patient is unconscious or otherwise incapable of consenting, and harm from a failure to treat is imminent and outweighs any harm threatened by the proposed treatment.”117 The exception ordinarily would include treatment for pain in such circumstances.

The emergency exception is actually quite narrow. It certainly does not give the ED carte blanche to treat every ED patient without consent. It only applies where the patient’s condition is urgent and the time required for consent would put the patient at serious risk of death or severe injury.118 In the case of the incapacitated patient, the nurse or doctor should secure the consent of a family member or other surrogate where possible without serious harm to the patient.

In litigation alleging emergency treatment without consent, several courts have concluded that consent for emergency treatment is implied by the patient’s coming to the ED.119 This implied consent does not extend to situations where the physician knows that the patient objects to treatment or particular interventions, however;120 and the notion of implied consent should not be relied upon too broadly. Quite frequently, an ED patient will sign a general consent form. Even with the general consent, the care provider should continue to inform the patient concerning his or her treatment; and a more specific consent should be sought for any procedure or medication with serious risks. For example, procedural sedation presents risks of damage to the central nervous system and depression of cardiac and respiratory functions. ACEP policy states that implied consent may be acceptable where the patient is unable to understand the necessary information due to altered mental status or severe pain and anxiety.121 Otherwise, separate consent to sedation is recommended.122

The key components to informed consent are that the patient is able to understand what options exists as well as the consequences of choosing one over another and is able to evaluate the costs and benefits of these consequences by relating them to a framework of values and priorities.123 One of the most serious problems regarding informed consent in the ED is the difficulty in ascertaining whether the patient is incapacitated. The stress and duress of an emergency condition, especially one associated with severe pain, may compromise the ability of the patient to consent; but the patient will not be legally incapacitated. The same judgment call is required for patients whose mental state is impaired by abuse of drugs or alcohol. Of course, the characteristics of the relationship between emergency doctor and patient, as described earlier, make a judgment about this individual’s preferences and values quite difficult. In that regard, the latitude that courts have allowed emergency physicians in the face of challenges to a lack of informed consent reflects this situation.

ED doctors may also be concerned that opioid analgesia will incapacitate the patient and make it impossible for that patient to consent to necessary treatments. In fact, severe pain may interfere with the patient’s ability to receive information and make rational risk assessments, although the patient will not be legally incapacitated, and doctors should not withhold opioid pain medication entirely for concern over incapacitating the patient.124

In regard to any medication that may impair judgment, alertness, or physical capacity, including pain medication, the physician must inform the patient clearly and accurately of these limitations prior to discharge. The physician, for example, should warn the patient specifically if the medication could interfere with driving or other similar activity and document this warning. Inadequate warnings have triggered physician liability in some cases.125 At least one study of pre-
scribing upon discharge from the ED cautions ED doctors to intensify efforts in this regard because 7% of patients in that study admitted to driving while taking narcotics within 7-14 days of discharge. Patients may choose among different options, with differing levels of effectiveness and adverse effects, for treatment of pain. Some patients may forego the most effective pain relief if it will compromise other goals. Physicians and nurses need to educate their patients so that the patient is not making this decision based on inaccurate assumptions about the potential for sedation or addiction.

The fact that EMTALA does not clearly mandate treatment for pain does not mean that such treatment is not otherwise legally required.

The question of authority arises in two ways: is the professional authorized to provide the intervention under hospital policy, and is the professional authorized to do so within his or her scope of practice under state law? The scope of practice of non-physician health care professionals varies widely among the states and significantly among individual facilities. Scope of practice is significant. If a professional exceeds his or her statutory scope of practice, it is likely, absent exculpatory circumstances, that this action will be viewed as negligence per se without further proof of the standard of care; however, some states treat this situation only as evidence, but not conclusive evidence, of negligence.

Limitations on scope of practice, whether established by statute, custom, or the specific facility have a direct impact upon the treatment of emergency patients in pain. The expressed purpose of such limitations is to assure quality of care, and so they are intended to improve the care of patients and may, in fact, do so. These limitations also directly effect access to treatment through, for example, limiting prescribing authority or requiring direct physician supervision of the non-physician professional. One of the areas of particular concern in the context of emergency treatment is the significant delay in providing treatment for the relief of severe, acute pain, as discussed earlier. Limitations on the scope of practice of emergency medical service professionals need to be examined in this context.

A related but distinct legal issue arises in the context of procedural sedation and other similar interventions. Although the procedure may be within the scope of practice allowed the professional under state licensure, the professional must also be competent by virtue of education, training and experience of performing the procedure. For example, a physician license is not limited to a particular range of medical practice, but not all physicians are competent to perform procedural sedation. ACEP policy asserts that all emergency physicians should be capable and competent in performing procedural sedation and that an anesthesiologist is not ordinarily required.

Emergency Medical Treatment and Labor Act
If there is a 500-pound gorilla in the ED, it is the federal Emergency Medical Treatment and Labor Act (EMTALA). EMTALA requires that a hospital receiving Medicare and operating an ED provide to any individual who “comes to” the emergency department with a request for aid an “appropriate medical screen-
ing examination...to determine whether or not an emergency medical condition exists.” If the hospital determines that the individual has “an emergency medical condition,” the hospital must provide medical treatment to “stabilize” the condition or, alternatively, arrange for transfer through appropriate means if the patient requests transfer or if the physician (or another authorized person) certifies that the “medical benefits” of transfer outweigh the increased risks of transfer.130

ED policies and practices are organized toward documenting compliance with the Act. It is probably the most significant legal concern that EDs and emergency medicine doctors have. If the Act were to clearly establish a legal duty for pain relief, it would be likely to have a very significant effect. Unfortunately, the answer to whether EMTALA requires treatment for pain is not entirely clear.

**Pain Assessment in the “Appropriate Medical Screening Examination”**

The courts have consistently held that the EMTALA requirement for an “appropriate medical screening examination” to determine whether the patient has an emergency medical condition requires no more than that the hospital screen each and every ED patient in the manner of the hospital's usual policy, custom and practice.131 The courts have refused to apply general professional standards of care to the screening requirement. Thus, the courts are unlikely to adopt the policies on pain management from organizations such as ACEP and JCAHO, discussed earlier, as the legal standards for compliance with EMTALA's medical screening requirement. With the implementation of the JCAHO standards on pain assessment, however, each accredited hospital now probably includes assessment for pain within their usual and customary initial and ongoing assessment and medical screening exam process. Once the hospital adopts this as practice or policy, pain assessment becomes a required element of the appropriate medical examination required under EMTALA. In addition, because the Act specifically recognizes “severe pain” as a symptom of an emergency medical condition, it may be argued that pain assessment is an essential part of any screening. Finally, courts may in the very rare case hold that a hospital's standard policies, procedures, and practices are so deficient as to amount to no medical screening at all. An evaluation of the patient’s report of pain is an essential diagnostic tool, and a failure to assess pain is likely to meet this extreme standard.

If EMTALA requires pain assessment at all, it is clear that the pain assessment is required at various points during the patient’s care in the ED and particularly upon discharge. The Interpretive Guidelines, issued by the Centers for Medicare and Medicaid Services for the reviewers who test compliance with or investigate violations of EMTALA, state that CMS believes that a medical screening examination “is an ongoing process;” that “the record must reflect continued monitoring according to the patient’s needs;” and that there “should be evidence of this evaluation prior to discharge or transfer.”132

**Pain and the Emergency Medical Condition**

The statute defines “emergency medical condition” as “a medical condition manifesting itself by acute symptoms of sufficient severity (including severe pain) such that the absence of immediate medical attention could reasonably be expected to result in placing the health of the individual...in serious jeopardy; serious impairment to bodily functions; or serious dysfunction of any bodily organ or part.”133 The statute references severe pain, not as an emergency medical condition itself, but rather as a symptom of an emergency medical condition. Despite the explicit reference to pain as a symptom of an emergency medical condition and despite the likelihood that the hospital's customary medical screening includes a pain assessment, it is not clear that EMTALA requires treatment for pain. EMTALA appears, then, to adopt the traditional, though now dated, perspective of emergency medicine that pain is merely a symptom.

**Stabilization and the Relief of Pain**

The statutory definition seems to anticipate that a patient may have the symptom of severe pain – a manifestation of an emergency medical condition – but not actually have an emergency medical condition. The EMTALA treatment requirement is limited to that treatment required to “stabilize” the patient. Stabilization is defined as providing “such medical treatment of the condition as may be necessary to assure...that no material deterioration of the condition is likely to result from or occur during [transfer].”134 Unless pain will result in a material deterioration of the patient's emergency medical condition, treatment for the pain itself is not required under EMTALA.

Thus, EMTALA does not ordinarily require that the ED have the patient’s pain managed prior to discharge or transfer unless the pain will cause the patient’s medical condition, as defined in the Act, to deteriorate as a result. This conclusion may be limited, however. Under EMTALA, the adequacy of the medical treatment required to stabilize the patient is measured against professional standards of care, not the hospital's own practices. EMTALA incorporates a malpractice standard in reviewing the adequacy of treatment of persons with an emergency medical condition. As medical practice begins to view interventions to relieve pain as essential to...
minimally adequate care, the emerging standards may infiltrate EMTALA cases either through the malpractice standard for stabilization or because of new understandings and evidence of what constitutes a “material deterioration” of a patient’s emergent condition and how unrelieved pain can result in such deterioration. Moreover, where the emergency medical condition is mental or emotional, conditions also within the EMTALA obligation, unrelieved pain itself may be a cause of material deterioration in the patient’s medical condition.

In most EMTALA cases litigated, in fact, the complaint is the failure to diagnose and treat life-threatening medical conditions such as myocardial infarction, either because the screening examination was inadequate or because treatment was inadequate. In the typical case, the pain was addressed through medication, but the underlying condition was not.

The impact of EMTALA, if extended to encompass a duty to provide adequate pain management as a part of the duty to stabilize, would be tremendous. Emergency departments orient their documentation, and thus their procedures, toward EMTALA compliance. Nevertheless, the fact that EMTALA does not clearly mandate treatment for pain does not mean that such treatment is not otherwise legally required. Medical malpractice and other tort claims, such as those described above, will still apply.

State and Federal Regulation of Prescribing Practices

Emergency physicians, nurses, and other professional and paraprofessional health care workers are subject to regulation through state licensure and through other state regulations involving the health and safety of patients and health care workers. Work in the ED is regulated by several federal agencies, including the Occupational Health and Safety Administration, the Food and Drug Administration, and the Centers for Medicare and Medicaid Services, among others.

Prescribing of medications that are listed as controlled substances in the schedules of the federal Controlled Substances Act (CSA) is regulated at both the state and federal levels. Physicians’ fear of regulatory scrutiny and intervention on the part of the state bureau of narcotics, the state medical licensure board, and the federal Drug Enforcement Administration (DEA) is a substantial barrier to access to effective pain relief for patients. The fear of providing controlled substances to patients with no medical need for the drugs also appears to be a substantial fear among ED physicians. It is not clear whether this behavior on the part of emergency physicians is attributable to a fear of legal sanction or a more culturally embedded concern for being tricked by duplicitous individuals posing as patients, as discussed earlier.

The public policy concerns underlying the Controlled Substances Act and licensure sanctions for prescribing practices are the risk of addiction and the diversion of certain medications. The public policy challenge in implementing both the CSA and state standards con-
“drug seeking” and “pain relief seeking” behaviors; and this confusion can penalize particular patient groups.

The emergency physician should engage in reasonable practices to assure that prescriptions for controlled substances meet current standards, such as those offered by the Federation of State Medical Boards, but adjusted to the practice of emergency medicine.

In recent years, the concept of “balance” has been used to provide a common meeting ground for those concerned with diversion and abuse of prescription drugs and those concerned with improving the care of patients in pain. The term, as commonly used in this context refers to a fundamental principle that government policies to prevent misuse of controlled substances should not interfere in their essential uses for the relief of pain. Implicit in the concept is recognition that both unrelieved pain and addiction are public health issues.

Thus, “balance” is a regulatory goal. It is not a principle that translates directly to clinical practice with an individual patient. Physicians surely must assess the benefits and risks of any medication for the individual. That exercise – by the physician and the patient together – involves a balancing function; but it is a balancing function of the risks and benefits to this particular patient, individuated by what is actually known about specific risks of addiction for particular groups of patients. This clinical assessment focuses on this patient’s particular needs, risks and overall best interests. The physician should not balance the general risk of abuse in the general population against this particular patient’s best interests.

State medical boards have made significant progress in adjusting their requirements for disciplinary actions to better reflect emerging standards of care for the treatment of patients in pain. The Federation of State Medical Boards issued guidelines for medical boards in 1998, and revised them in 2004. These guidelines, adopted by many states, clearly state that fostering effective pain relief is a goal of the regulatory process; that physician prescribing will not be judged by volume or chronicity alone, but rather by outcomes for the patients; and that the physician has an obligation to perform and document a physical examination of the patient and a care plan that includes appropriate follow-up. At least 23 state legislatures have enacted “intractable pain statutes” to further affirm the importance of treating pain, and to set out some guidance for appropriate regulatory oversight of prescribing practices. In fact, some medical boards have taken disciplinary action against physicians who have neglected their patients in pain.

At the same time as state regulatory standards and enforcement efforts are accommodating a goal of improving quality of care for patients in pain while attending to their obligations to protect against addiction and diversion, the federal government has intensified its efforts against the prescribing of controlled substances for pain management, and have engaged in a strategy of high profile arrests and prosecutions of physicians. In addition, the DEA has parted ways with the approach developed in the majority of states.

The purpose of the Controlled Substances Act, enforced by the DEA, is to control illegitimate distribution of controlled substances without interfering with legitimate medical and scientific practices. The tension between the states and the DEA on what qualifies as legitimate medical practice is growing in this issue and in others. The ideal, however, is that the physician be guided by the same or at least consistent standards as between federal and state regulators.

In recognition of the establishment of new practice standards in the states and the inadequacy of pain management in the U.S., the DEA issued a statement in 2001 advocating a balanced regulatory policy for prescription pain medications that would account both for concerns over addiction and diversion and concerns over pain management. In this statement, joined by 21 national organizations, the DEA recognized the regulatory balance: “We want a balanced approach that addresses the abuse problem without keeping patients from getting the care that they need and deserve.” The DEA took another pragmatic step toward achieving a more balanced approach to its enforcement efforts in 2003, when the agency issued a “Frequently Asked Questions” document (the FAQs). The approach to oversight of prescribing practices for pain management taken in the FAQs was consistent with the model guidelines published earlier by the Federation of State Medical Boards. This development brought state and federal efforts into harmony, allowing physicians to practice in a more predictable environment. The harmonization was particularly welcome because several states had become more interested in penalizing physicians for reckless disregard of pain through disciplinary actions and private parties had brought two very high profile personal injury cases, as discussed above.

The FAQs provided educational information to medical practitioners through a series of questions and answers about the appropriate use of opioids in the treatment of pain. The FAQs addressed the definition of pain and its treatment; how opioids work and what patients need to know; the risks in the medical use of opioid analgesics; and legal and regulatory considerations, including under what circumstances the DEA would be likely to decide to investigate and what medical professionals needed to do to comply with state and federal law.
Subsequent to their publication, the FAQs were immediately embraced by the professions that were pursuing ways to address the inadequate treatment of pain. In an effort to dispel the fear of legal sanction that was impeding appropriate prescribing, the FAQs were held out as an indication that physicians who comply with particular standards of patient care could do so without fear of investigation or sanction.\textsuperscript{149} Even though emergency physicians had not been the particular targets of DEA action, the literature in emergency medicine also recognized the significance of the positive changes in the legal environment on a federal and state level.\textsuperscript{150} One such article, for example, used the FAQs to encourage emergency department professionals to abandon their fear of legal risks and “appreciate the greater protections offered...when operating by acceptable medical standards.”\textsuperscript{151}

Soon after the DEA issued the FAQs, however, the agency’s commitment to the “balanced” approach began to crumble. The retrenchment began in 2003, with the release of the statement “The Myth of the Chilling Effect” on the DEA’s web site.\textsuperscript{152} This statement identifies the mission of the DEA: “to prevent, detect and investigate the diversion of legitimately manufactured controlled substances.” The statement does not specifically affirm the importance of the treatment of pain as did the 2001 joint statement and the FAQs. The statement asserts that “doctors operating within the bounds of accepted medical practice have nothing to fear from the DEA,” but it does not give specific guidance as to the “bounds of accepted medical practice.” The statement simply provides statistics on DEA’s enforcement efforts, noting that the agency had “pursued sanctions against less than one tenth of one percent of the registered doctors” since 1999.

What was shaken by the posting of “The Myth of the Chilling Effect” was completely disassembled by the retraction of the FAQs by the DEA in November 2004. An interim policy statement (IPS) published by the DEA in the Federal Register announced the withdrawal of the FAQs,\textsuperscript{153} citing “misstatements” in the FAQs. The IPS clearly rejects the approach to oversight that had been adopted by the Federation of State Medical Boards and by many states. With the withdrawal of the FAQs and the substantive statements made in the IPS, the DEA has taken federal regulation and oversight for prescribing for pain in a direction that is the opposite of that taken by the majority of the states. In a letter to the DEA after the retraction of the FAQs, the National Association of Attorneys General expressed concern that “the state and federal policies are diverging with respect to the relative emphasis on ensuring the availability of prescription pain medications to those who need them.”\textsuperscript{154}

Although the current regulatory environment, as played out by the state medical boards and the DEA, is a difficult one for doctors treating patients in pain, and particularly chronic pain patients, the emergency department physician is somewhat insulated from the fray. Federal enforcement efforts have and probably will continue to target the office-based medical practice rather than the hospital-based emergency medicine practice. Still, even if the emergency doctor is not at particular risk of enforcement activity, chronic pain patients who are ill-served by the current regulatory environment are likely to show up at the doors of the emergency department.

**Conclusion**

We know little of what we need to know to improve the treatment of patients in pain who are seeking care in the emergency departments in the U.S. That we have reasons to improve that care is clear. Recognized ethical duties; enforceable legal obligations; and human compassion and empathy all drive us toward that goal. In the case of the emergency department, the seriousness of untreated pain may be underestimated if it is viewed as merely a temporary experience. Enough research exists, however, for us to be able to argue that the impact is long term.

Efforts at improving care nearly always begin with trying to discover the reason for the failure of care—discovering the “root cause,” so to speak. With neglect of pain generally, we still often deal with questions: Does information change practice? Will a change in legal enforcement policies change practice? With emergency medicine, we may have even less knowledge about the reasons physicians behave the way they do. Studying emergency medicine in context, however, gives us the opportunity to look at now familiar problems in what is a very different medical culture than either the office-based or the palliative care settings, and one which struggles with uncertainty, unfamiliarity, and subjectivity.

Further research is absolutely critical. The research needs to focus on the issues that lead emergency physicians to withhold interventions that could help patients as well as on the basic clinical research on the effectiveness and safety of certain interventions. Conducting research in the context of emergency medical care is very challenging, but it is worth it.

**Acknowledgements**

The author gratefully acknowledges the support of the Mayday Fund for this research as well as the research assistance of Carine M. Boyle.

**References**

1. See, e.g., House Resolution 1020, introduced on March 1, 2005, “to


36. Id.

37. E. D. Fosnocht, E. R. Swanson, P. Bossart, “Patient Expectations for Pain Medication Delivery,” American Journal of Emergency Medicine 19, no. 5 (2001): 399-402, reporting that patients’ expected to receive their first administration of pain medication on average within 23 minutes of arrival while the actual time to administration was 78 minutes.


45. Id.


PAIN MANAGEMENT IN THE EMERGENCY DEPARTMENT • WINTER 2005 759


135. 21 U.S.C. s 801 et seq.


143. Id., at 17.

144. 11 BNA Health Law Reporter 1222.

145. Gonzales v. Raich, 125 S.Ct. 2195 (2005).

146. Gonzales v. Raich, 125 S.Ct. 2195 (2005); Gonzales v. Oregon, 368 F.3d 1118.


148. Prescription Pain Medications: Frequently Asked Questions and

Sandra H. Johnson


Pain is the single most common reason patients seek care in the emergency department.\textsuperscript{1} Given the prevalence of pain as a presenting complaint, one might expect emergency physicians to assign its treatment a high priority; however, pain is often seemingly invisible to the emergency physician. Multiple research studies have documented that the undertreatment of pain, or oligoanalgesia, is a frequent occurrence.\textsuperscript{2} Pain that is not acknowledged and managed appropriately causes dissatisfaction with medical care, hostility toward the physician, unscheduled returns to the emergency department, delayed return to full function, and potentially, an increased risk of litigation. Failure to recognize and treat pain may result in anxiety, depression, sleep disturbances, increased oxygen demands with the potential for end organ ischemia, and decreased movement with an increased risk of venous thrombosis.

Given this state affairs, we should examine the barriers that serve to block the adequate recognition and treatment of pain in emergency departments, as well as other healthcare delivery settings. One of these barriers is the physician’s fear of being “duped” by patients who fabricate pain symptoms in order to obtain controlled substances for recreational use or diversion. This article will focus on the problem of substance abuse, and the perceptions of healthcare providers regarding substance abuse, as they relate to patients who present to the emergency department with complaints of pain.

**The Prevalence of Pain**

Pain is a near universal human experience. Acute pain can be defined in terms of duration: characteristically it is of recent onset and lasts no more than a few days to several weeks. It usually occurs in response to tissue injury and disappears when the injury heals. Acute pain serves an adaptive purpose in that it is associated with protective reflexes, such as withdrawal responses to remove a limb from danger, or muscle spasms that serve to immobilize an extremity; however, some responses associated with acute pain may be maladaptive, leading to impaired immune responses, elevated myocardial oxygen demands, hypercoagulation, and atelectasis.

While less common, chronic pain affects approximately one third of the U.S. population annually.\textsuperscript{3} Domestic and international survey studies have reported chronic pain prevalence rates as high as forty percent.\textsuperscript{4} Traditionally, chronic pain has been defined as pain

---

**Chronic Pain and Aberrant Drug-Related Behavior in the Emergency Department**

*Knox H. Todd*

---

**Knox H. Todd, M.D.,** is Professor of Emergency Medicine at the Albert Einstein College of Medicine and Director of the Pain and Emergency Medicine Institute at Beth Israel Medical Center in New York.
Of particular importance to the emergency physician’s assessment of pain, and as one potential explanation for the phenomenon of oligoanalgesia, acute and chronic pain can be associated with markedly different pain-related behaviors.

Of particular importance to the emergency physician’s assessment of pain, and as one potential explanation for the phenomenon of oligoanalgesia, acute and chronic pain can be associated with markedly different pain-related behaviors. While acute pain is usually associated with objective signs of sympathetic nervous system activation and overt signs of physical suffering, patients with chronic pain may not exhibit such typical behaviors and signs of autonomic nervous system overactivity. This disparity between observed patient behaviors and physician expectations of such behaviors in the setting of chronic pain may lead to inaccurate determinations of pain intensity and ultimately, the undertreatment of pain.

Several studies have attempted to define the prevalence of pain in emergency department settings. Johnston, et al., conducted a prospective study to determine the incidence and severity of pain among patients presenting to non-critical treatment area within the emergency departments of two urban hospitals in Canada. Fifty-eight percent of adults and forty-seven percent of children reported pain on emergency department arrival. For approximately one-half of both groups reporting pain, the intensity of this pain was considered moderate to severe. At the time of discharge, one-third of both groups continued to experience pain of moderate to severe intensity. In fact, eleven percent of children and adults in this study actually reported clinically important increases in pain intensity during the emergency department stay.

A second prospective study, conducted by Tanabe and Buschmann, found that among adults treated at one Chicago emergency department, fully seventy-eight percent presented with a chief complaint related to pain. Of these patients, only fifty-eight percent received analgesics or nonpharmacologic interventions to treat pain. For patients receiving analgesics, an average of seventy-four minutes elapsed from the time of arrival to the time of treatment. Only fifteen percent of patients were treated with opioids, despite high levels of pain intensity. Interestingly, sixteen percent of patients with pain in this study indicated that they would have refused analgesics had they been offered. The principal reported reason for their refusal was the fear of addiction resulting from opioid exposure, even when opioids were indicated for the treatment of pain.

In 2002, Cordell, et al., reported an analysis of secondary data from an urban, tertiary-care emergency department using explicit data abstraction rules to determine the prevalence of pain and to assign painful conditions into standard categories. With inclusion of all age groups, they found evidence of pain in sixty-one percent of patients. Pain was the chief complaint for fifty-two percent of patient visits. After excluding patients less than five years of age for whom chart reviews are obviously less reliable, almost seventy percent of patient encounters were determined to involve pain complaints.

While the high prevalence of pain among emergency department patients is well documented, the underlying conditions responsible for pain in this population are less well characterized. In Cordell’s retrospective study, eleven percent of patients presenting to the emergency department were judged to be suffering from pain that was chronic in nature. In a recent prospective multicenter study conducted in the U. S. and Canada, forty-four percent of ultimately discharged patients presenting to the emergency department with pain reported underlying chronic pain syndromes. In one-half of these cases, the emergency department visit was prompted by an exacerbation of this chronic pain condition. Importantly, patients with chronic pain reported three to four times the number of annual physician visits when compared to those without chronic pain. Median and mean durations of symptoms for those re-
porting chronic pain syndromes were twenty-four and fifty-two months, respectively. For physicians who view themselves as experts in the management of acute medical and surgical emergencies, chronic pain may represent a less familiar condition with which to contend.

The Prevalence of Substance Abuse

Definitions

In discussing issues of chemical dependency and aberrant behaviors related to opioid use, a valid system of nomenclature is necessary for clear communication and measurement. Historically, the meaning of different terms has changed, particularly in light of the increased use of chronic opioid therapy for malignant and non-malignant chronic pain conditions. In treating pain in this population of patients with chronic opioids, confusion over the concepts of physical dependence, tolerance, addiction, and pseudoaddiction may constitute a barrier to understanding and to appropriate treatment. These phenomena are discrete and standard definitions may be helpful in caring for such patients. Currently accepted definitions of these terms are as follows:9

Addiction is a primary, chronic, neurobiologic disease, with genetic, psychosocial, and environmental factors influencing its development and manifestations. It is characterized by behaviors that include one or more of the following: impaired control over drug use, compulsive use, continued use despite harm, and craving.

Physical dependence is a state of adaptation that often includes tolerance and is manifested by a drug class specific withdrawal syndrome that can be produced by abrupt cessation, rapid dose reduction, decreasing blood level of the drug, and/or administration of an antagonist.

Tolerance is a state of adaptation in which exposure to a drug induces changes that result in a diminution of one or more of the drug's effects over time.

Pseudoaddiction is a term which has been used to describe patient behaviors that may occur when pain is undertreated. Patients with unrelieved pain may become focused on obtaining medications, may “clock watch,” and may otherwise seem inappropriately “drug seeking.” Even such behaviors as illicit drug use and deception can occur in the patient's efforts to obtain relief. Pseudoaddiction can be distinguished from true addiction in that the behaviors resolve when pain is effectively treated.

The term, “substance abuse” is particularly problematic and resistant to precise definition. The American Psychiatric Association has defined substance abuse as a maladaptive pattern of drug use associated with some manifest harm to the user or others.10 Other groups using consensus methodology have defined abuse as any use considered to be outside of socially accepted norms.13 Determining the bounds of “socially accepted norms” within the broad range of social strata treated within any emergency department is a difficult task. Physicians may believe that they “know abuse when they see it” and its identification may be influenced by subjective judgments that may, or may not, correspond to socially accepted norms for the index patient’s particular social group. Often the term, “substance misuse” is applied to behaviors that are not perceived as particularly extreme, e.g., taking opioid analgesics to relieve symptoms other than pain such as anxiety or boredom.

The difficulty in determining whether a given set of behaviors fall within accepted definitions of substance use, misuse, or abuse has important implications outside the clinical realm. Physicians may prescribe controlled substances for the treatment of pain while patients may use these drugs to treat a broad range of symptoms with varying degrees of relatedness to underlying pain syndromes and may, in fact, use drugs in a manner totally unrelated to the physicians’ intent, i.e., to obtain euphoric, rather than analgesic, effects. Given the unclear distinctions between use, misuse, and abuse, and a regulatory climate in which practitioners prescribing patterns are increasingly scrutinized, emergency physicians are understandably reluctant to prescribe controlled substances to patients with whom they expect to have only a transitory relationship.

Using any definition, substance abuse is a highly prevalent problem. The National Survey on Drug Use and Health (formerly the National Household Survey on Drug Abuse) reports that in 2003, an estimated 19.5 million Americans, or 8.2 percent of the population aged twelve or older, used an illicit drug during the month prior to the survey interview. Illicit drugs included marijuana, cocaine, heroin, hallucinogens, inhalants, and nonmedical use of prescription-type pain relievers, tranquilizers, stimulants, and sedatives.12 Importantly, the survey documents an increase in the lifetime reported nonmedical use of pain relievers between 2002 and 2003, from 29.6 million to 31.2 million persons. To be considered “nonmedical” use, the respondent had to take drugs not prescribed for them or take them only for the “experience or feeling” they caused. Specific analgesics showing statistically significant increases in lifetime use were (in order by magnitude): Vicodin, Lortab, or Lorcet; Percocet, Percodan, or Tylox; Hydrocodone; OxyContin; Methadone; and Tramadol.
In contrast to the prominence of emergency department-based data collection systems in efforts to monitor deleterious outcomes associated with substance abuse, relatively few studies have systematically assessed substance abuse prevalence and treatment needs in the emergency department population. As an example, the Drug Abuse Warning Network (DAWN) is a federally financed, public health surveillance system that monitors drug-related emergency department visits and drug-related deaths investigated by medical examiners and coroners. This reporting system involves hundreds of hospital emergency departments throughout the U.S. and provides valuable data with which to monitor drug abuse trends. In contrast to this large monitoring research enterprise, relatively little focus has been given to use of the emergency department as a setting in which to intervene in substance abuse problems.

In 1997, Soderstrom, et al., assessed the prevalence of psychoactive substance use disorders in a large, unselected group of seriously injured patients treated at a Level I trauma center in Baltimore, using standardized diagnostic interviews and explicit criteria. Psychoactive substance use disorders were diagnosed using the Structured Clinical Interview (SCID), an instrument based on the Diagnostic And Statistical Manual of Mental Disorders, Revised Third Edition (DSM-III-R). Of 1118 patients consenting to the study, more than half had one or more lifetime abuse or dependence psychoactive substance use disorders, and eighteen percent were currently considered dependent on drugs other than alcohol.

In 1996, Rockett, et al., used direct interviews to ascertain unmet substance abuse treatment needs in a statewide probability sample survey of adults presenting to seven Tennessee emergency departments. While only one percent of emergency department medical records indicated a diagnosis of alcohol or drug-related problems, as many as twenty-seven percent of patients were determined by the researchers to need substance abuse treatment on the basis of explicitly defined case definitions. Less than ten percent of patients that were ultimately determined to need substance abuse treatment in this study were actually receiving such care. Thirty-two percent of all patients in this study had a positive saliva or urine assay for psychoactive drugs and nine percent screened positive for opioid use. Unmet substance abuse treatment needs varied directly with the frequency of emergency department visits and inversely with patient age.

A subsequent study by Rockett, et al., examined the association between unmet substance abuse treatment needs in the emergency department and excess utilization of health services in order to estimate the health care costs savings that might result from effective emergency department-based substance abuse treatment interventions. The researchers estimated that patients with unmet substance abuse treatment needs accounted for an estimated 777 million dollars in extra hospital charges for Tennessee, or 1,568 dollars per emergency department patient when compared to those without substance abuse treatment needs. They suggested that the costs of emergency department-based screening and intervention efforts targeted to substance abuse disorders would be more than offset by savings from decreased health care utilization and that these programs were likely to be highly cost-effective if implemented.

### The Problem of “Drug-Seeking Behavior”

The preceding review makes clear the high prevalence of both pain and substance abuse disorders in the emergency department. Although acute and chronic pain is far more common than substance abuse disorders, it is inevitable that emergency physicians will frequently encounter patients presenting with both pain and substance abuse disorders. Professional discussions of pain treatment in the emergency department frequently center on concerns of being duped by such patients who fabricate painful symptoms in order to obtain opioids, so-called “drug-seeking behavior.” “Drug-seeking behaviors” may represent an entirely appropriate response by those with chronic pain who are routinely undertreated by the medical profession and for whom comprehensive pain treatment centers are in short supply. Although the term, “drug seeking behavior,” is poorly defined, it is used in the emergency medicine literature and will be used in this article, with acknowledgement of its imprecision.

Only a limited amount of emergency medicine research has addressed this problematic issue. In 1990, Zechnich and Hedges attempted to measure community-wide use of emergency department services by patients at high risk for drug seeking behavior. In this retrospective, observational study patients were categorized as exhibiting drug seeking behavior if they sought care at a university hospital in Portland, Oregon for a specific pain-related diagnosis (i.e., ureteral colic, toothache, back pain, abdominal pain or headache) and were either independently identified on at least one other local hospital’s “patient alert” list or suffered a drug-related death during the year in question. After identifying thirty-three such patients, they determined the frequency of their emergency department visits at each of seven local hospitals and conducted detailed chart reviews of their visits at three of these hospitals. The patients identified as drug seeking were generally young and one-half of drug seekers were female. The
latter is a surprising finding; given that substance abuse disorders are more than twice as common among males.\(^9\) This suggests that drug seeking behaviors are exhibited (or identified) more commonly among female emergency department patients with substance abuse problems than among males.

The thirty-three patients visited emergency departments, urgent care clinics, or were hospitalized a total of 379 times over the study period, for an average of 12.6 visits per person annually. Interestingly, although chart reviews identified seventeen patients who were told that he or she “would receive no further narcotics” at a given facility, these patients subsequently received controlled substances from another hospital in ninety-three percent of cases and from even the same facility in seventy-one percent. The authors suggested that information sharing between hospitals could help to identify drug seeking patients and promote more consistent community-wide care and appropriate substance abuse interventions.

The maintenance of lists that include the names and medical information for patients frequently seen in the emergency department is thought to be a common practice. In a mail survey conducted in 1995, Graber, et al., described the use of what were referred to as “problem patient files” in the state of Iowa. Fifty-eight percent of emergency department medical directors acknowledged the use of such files and responded that the files were consulted an average of 2.6 times per week. Calls between emergency departments either seeking or responding to requests for information about patients listed in these files were estimated to occur twenty-three and twenty times per year respectively. Rarely were explicit policies established for limiting access to these files and information was added to the records in an informal fashion.\(^20\)

In 2000, Pope, et al., from Vancouver described a case management program for frequent visitors to their inner-city tertiary care emergency department serving a large number of patients with multiple psychosocial problems, including homelessness and substance abuse.\(^21\) Of twenty-four patients described in this study, five were said to exhibit drug seeking behavior, and eight patients suffered from alcohol and drug abuse, personality disorders, and chronic pain. These twenty-four patients accounted for a staggering 616 visits annually (median 26.5 visits per year). After the implementation of individualized chronic care plans that included social work interventions at the time of the visits, emergency department use by this group of superutilizers dropped to a median of 6.5 visits per person per year.

In 2003, a publication by Geiderman discussed ethical, legal, and regulatory considerations surrounding the use of what were termed “habitual patient files.”\(^22\) The article acknowledged the common and informal use of such files, and set forth standards intended to promote the development of formal policies and procedures to govern their use. The author noted that such files have never been demonstrated to be effective in either reducing emergency department use by drug seeking patients or in altering care patterns and suggested the need for a research program to explore the impact of their use. Finally, the author called for a coordinated and comprehensive program of physician education to promote the identification and treatment of emergency department patients with substance abuse disorders.

**Pain and Substance Abuse: A Balanced Perspective**

In managing pain, emergency physicians are responsible for beneficence as well as nonmaleficence. We must treat pain and ameliorate suffering while minimizing the extent to which our treatment strategies enable substance abuse by our patients. For the vast majority of patients presenting with acute monophasic pain, whether from trauma, acute medical illness, or procedures performed in the emergency department, there is little danger of enabling substance abuse and a great deal of room for improvement in the quality of analgesic practices. Multiple published studies have documented the continued prevalence of oligoanalgesia among children and adults treated in our emergency departments.\(^23\)

For a small subset of emergency department patients, particularly for those presenting with chronic or recurrent pain syndromes, the physician may have legitimate concerns regarding an underlying substance abuse or related disorder. Our task is to balance the often unclear risk of fostering substance abuse, and even diversion, in this subset of patient with the well-known and well-documented risk of undertreating painful conditions.

To the extent we can clarify the nomenclature used to classify patients with pain and substance use disorders,
we can begin to identify more effective approaches to both problems. To aid in this effort, we will attempt to clarify various phenomena that have been lumped within the term, “drug seeking behavior.”

To begin, it must be said that “drug seeking behavior” is a term best abandoned by our profession. For the patient in pain, seeking an analgesic of proven effectiveness is the height of rationality. In contrast to the search for controlled substances, it is likely that the most common variety of drug seeking behavior is the well-documented and relentless quest by patients with self-limited viral upper respiratory infections (or parents of such patients) to obtain antibiotics. The medical profession has a long history of inappropriately prescribing such antibiotics, encouraging antibiotic resistance among common bacterial strains while risking antibiotic side-effects without a justifiable expectation of concomitant benefit.

The concern of physicians is that patients may seek controlled substances, particularly opioids and benzodiazepines, for reasons other than those strictly related to pain relief. Such actions are best termed “aberrant drug-related behaviors” as this term suggests that there is a broad range of behaviors that are more acceptable or less acceptable in the context of pain therapy. Although addiction is the most commonly assumed explanation for such aberrant behaviors, there is an extended differential diagnosis for such behaviors that the clinician should consider.

Although confirmatory research is lacking, expert consensus suggests that aberrant drug-related behaviors reflect a broad range of observed activities that are either more, or less, suggestive of an addiction disorder. (Table 1) Certainly, the presence of an obvious painful condition (e.g., appendicitis, fracture) should preempt concerns about illegitimate drug-seeking behaviors. At the other extreme, even behaviors that are clearly unacceptable do not necessarily indicate addiction or diversion. Hay and Passik have even reported one case of prescription forgery that was seemingly unrelated to addiction or criminal intent. The forgery occurred when the patient’s caregiver was leaving for vacation, prompting excess anxiety and fear of abandonment.24 Addiction is but one of many diagnoses that may lead to aberrant drug-related behaviors. (Table 2)

Given the high prevalence of chronic pain and the widespread unavailability of chronic pain management resources, particularly for populations served by the emergency department, pseudoaddiction is the most likely cause for a large proportion of drug-related behaviors deemed aberrant. In particular, patient reports of distress associated with unrelieved symptoms, aggressive complaining about the need for higher doses of analgesics, and unilateral dose escalation by the patient are suggestive of pseudoaddiction. Establishing the diagnosis of pseudoaddiction is particularly difficult if the patient has both pain and a comorbid substance use disorder; however, the two can obviously coexist. The signature of pseudoaddiction is that aberrant behaviors disappear when adequate analgesics are given to control pain.

The condition that best exemplifies the problem of emergency department-based pseudoaddiction is sickle cell disease. Vaso-occlusive pain crises are the most common reason for emergency department visits by patients with sickle cell disease and the genetics, molecular biology, and pathophysiology of this disease are relatively well understood. Although the management of sickle cell vaso-occlusive pain crises is viewed as challenging by emergency physicians, it has been a relatively neglected area of research investigation by the specialty.25 Despite the fact that almost all of the 75,000 annual hospitalizations for pain crises occur after emer-

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Spectrum of Aberrant Drug-Related Behaviors that Raise Concern about the Potential for Addiction.</strong></td>
</tr>
<tr>
<td><strong>Less suggestive of addiction:</strong></td>
</tr>
<tr>
<td>• Aggressive complaining about the need for more drug</td>
</tr>
<tr>
<td>• Drug hoarding during periods of reduced symptoms</td>
</tr>
<tr>
<td>• Requesting specific drugs</td>
</tr>
<tr>
<td>• Openly acquiring similar drugs from other medical sources</td>
</tr>
<tr>
<td>• Occasional unsanctioned dose escalation or other non-compliance</td>
</tr>
<tr>
<td>• Unapproved use of the drug to treat another symptom</td>
</tr>
<tr>
<td>• Reporting psychic effects not intended by the clinician</td>
</tr>
<tr>
<td>• Resistance to a change in therapy associated with “tolerable” adverse effects with expressions of anxiety related to the return of severe symptoms</td>
</tr>
<tr>
<td><strong>More suggestive of addiction:</strong></td>
</tr>
<tr>
<td>• Selling prescription drugs</td>
</tr>
<tr>
<td>• Prescription forgery</td>
</tr>
<tr>
<td>• Stealing or “borrowing” drugs from others</td>
</tr>
<tr>
<td>• Injecting oral formulations</td>
</tr>
<tr>
<td>• Obtaining prescription drugs from non-medical sources</td>
</tr>
<tr>
<td>• Concurrent abuse of alcohol or illicit drugs</td>
</tr>
<tr>
<td>• Repeated dose escalation or similar noncompliance despite multiple warnings</td>
</tr>
<tr>
<td>• Repeated visits to other clinicians or emergency rooms without informing prescriber</td>
</tr>
<tr>
<td>• Drug-related deterioration in function at work, in the family, or socially</td>
</tr>
<tr>
<td>• Repeated resistance to changes in therapy despite evidence of adverse drug effects</td>
</tr>
</tbody>
</table>

Despite our understanding of the sickle cell disease process, many health professionals are reluctant to prescribe adequate doses of opioids for these patients experiencing pain largely due to addiction concerns. In one survey study, fifty-three percent of emergency physicians were of the belief that more than twenty percent of patients with sickle cell disease were addicted to opioids, while only twenty-three percent of hematologists shared this belief. Also, in this survey thirty-five percent of hematologists reported that they followed pain management protocols when treating painful crises as compared to only seventeen percent of emergency physicians.

Nurses’ attitudes regarding the prevalence of addiction among this patient population are even more extreme, with sixty-three respondents reporting that addiction was prevalent. Thirty percent of nurses in this survey reported that they were hesitant to administer high-dose opioids for painful vaso-occlusive crises.

A hesitant approach to emergency department opioid administration in the setting of vaso-occlusive pain crises will predictably lead to continued pain, increased anticipation of pain, and increased patient anxiety. This experience may generate pain-avoidance manifestations by patients than are interpreted by physicians as aberrant drug-related behaviors. Eventually, larger doses of opioids may be administered to control pain that is spiraling out of control with resultant excessive sedation. This apparent sedation in the setting of a painful condition may reinforce the physician’s disbelief in the reality of his or her patient’s initial pain reports.

It has been demonstrated that this cycle of inadequate care can be broken by the institution of pain management protocols that emphasize continuous opioid infusions and sustained courses of orally administered controlled-release opioids. In 1992, Brookoff and Polomano reported the institution of such a structured analgesic regimen on hospital use by patients with sickle cell disease presenting to the emergency department of an inner-city university hospital in Philadelphia with remarkable results. After institution of the pain management protocol, the number of hospital admissions for sickle cell pain decreased by forty-four percent, the number of total inpatient days by fifty-seven percent, the hospital length of stay by twenty-three percent, and the number of emergency department visits by sixty-seven percent.

The authors asserted that these positive results were seen without a subset of patients being “chased away” from the hospital. Others have reported marked decreases in aberrant drug-related behaviors and the number of emergency department visits by patients with sickle cell disease after instituting long-term management of pain with chronic opioid therapies typically used to treat malignant pain.

Aside from considerations of pseudoaddiction, chronic pain is often accompanied by mood disorders and psychiatric comorbidities that complicate the management of these challenging patients. The presence of aberrant drug-related behaviors in patients with borderline personality disorders may represent an expression of fear and anger or an attempt to cope with chronic boredom. Patients may use opioids and alcohol in attempts to lessen symptoms of anxiety, panic disorder, depression, or insomnia. Emergency physicians often receive limited training in dealing with such disorders and the specialty’s deficiencies in dealing with such problems have been documented. Psychiatric consultation, if available, may be useful in both suggesting alternative causes for aberrant behaviors and tailoring the physician’s therapeutic approach to deal with these complicating factors.

For some patients, aberrant drug-related behaviors represent criminal intent to divert or sell controlled
substances. The prevalence of behaviors occasioned by such intent is unknown and it is likely that in many cases, multiple etiologies of aberrant behaviors coexist. Certainly, patients with active or past substance use disorders are at increased risk for injuries and illnesses that can lead to chronic pain (e.g., motor vehicle injury). Thus, the conditions listed in Table 2 are not mutually exclusive.

### Barriers to Improvement

Barriers to the treatment of pain are discussed by other authors in this issue. There is also a paucity of treatment guidelines and best practice standards for emergency department pain care, in part because of the lack of research in this area by emergency medicine investigators. Although the American College of Emergency Physicians has adopted a statement of general principles regarding pain management (Figure 1), the specialty lacks clearly articulated standards to drive pain care and healthcare systems do not include adequate mechanisms to ensure accountability for inferior practice.33

Given the concentration of patients with substance abuse disorders, the emergency department is an appropriate site for screening and intervention for both alcohol and drug problems; however, emergency physicians received limited training in recognition and appropriate interventions for such problems, and an air of pessimism characterizes physicians’ estimation of success for many substance abuse therapies. Translating our knowledge of therapeutic strategies into action against these disorders will require overcoming much clinical inertia.

Although federal regulators and state medical boards do not perceive emergency medicine as a specialty prone to inappropriate prescribing, and investigations of emergency physicians are rare, if not unheard of, many emergency physicians express fears of such scrutiny or sanctions related to prescribing or administering opioids. While this concern is often voiced, it seems likely that this fear represents concern about other, less obvious physician uncertainties related to pain management and substance abuse disorders. Emergency physicians may be concerned about being overburdened by the inherent difficulties of managing patients with complicated pain syndromes and coexisting substance abuse disorders.

In dealing with complex chronic pain patients, the emergency physician practicing in isolation may exhibit symptoms of despair and direct his or her anger toward the patient with pain, resulting in more alienation of patients who may have already been abandoned by other sectors of the healthcare system. This is particularly likely to happen in communities without multidisciplinary treatment centers for either substance abuse disorders or chronic pain and for those with inadequate healthcare insurance. Thus the patient with chronic pain joins the larger group of those with unmet healthcare needs that currently crowd our emergency departments. The hectic nature of emergency medicine practice often does not allow sufficient time for precisely characterizing patients with complex pain complaints and clinicians may lump legitimate pain behaviors with the ploys of those seeking opioids inappropriately. Both groups of patients may be ultimately mistrusted and treated with disdain.

Finally, the true prevalence of addiction and aberrant drug-related behaviors is unknown and unstudied. There is little research on risk factors for prescription drug abuse to guide the emergency physician. When the prevalence of such problems is overestimated, oligoanalgesia is the predictable result.

### Conclusion

Relieving pain and reducing suffering are primary responsibilities of emergency medicine and much can be done to improve the care of patients in pain. We have a concurrent duty to limit the personal and societal harm that can result from prescription drug abuse. Our specialty should continue to refine our approach to the
problem of pain and substance abuse and reduce the current large amount of variability in our practices. We should continue to more precisely define our own standards for excellence in pain practice and substance abuse interventions while promoting quality improvement initiatives to achieve these goals.

Acknowledgements
The author gratefully acknowledges the support of the Mayday Fund in the development of this article.

References
7. Cordell, supra note 1.
19. SAMHSA, supra note 12.
Background

The last decade has witnessed a dramatic shift in the perception of the safety of opioid prescribing. An initial analysis by the Pain and Policy Studies Group at the University of Wisconsin Medical School concluded that the use of opioids in the chronic pain population carried a low risk of abuse potential. They had studied the national utilization of five opioid analgesics used to treat severe pain along with a retrospective chart review of ED visits associated with abusive behaviors. From 1990 to 1996, there were increases in medical use of morphine (+59%), fentanyl (+1168%), oxycodone (+23%), and hydromorphone (+19%), and a decrease in the medical use of meperidine (-35%). During the same period, the total number of hospital emergency department admissions (resulting from drug abuse) per year due to opioid analgesics increased from 32,430 to 34,563 (6.6%), but the proportion of admissions for opioid abuse relative to total mentions for drug abuse decreased from 5.1% to 3.8%. The authors concluded that the trend of increasing medical use of opioid analgesics to treat pain did not appear to contribute to an increase in opioid analgesic abuse. But a subsequent epidemiologic study, the 2002 National Survey on Drug Use and Health, revealed that the number of people using prescription opioids for non-medical purposes (defined as use of prescription-type drugs not prescribed for the survey respondent by a physician or used only for the experience or psychomimetic feeling) increased significantly after 1996, the final year of data collection of the aforementioned Pain and Policy Studies Group report (Fig 1). In fact, by 2002, the prevalence of prescription opioid abuse surpassed that of cocaine and heroin and became second only to marijuana in terms of past year illicit drug dependence or abuse (Fig 2).

In the recent past, high-visibility press coverage of prescription opioid addiction has cast a bright light on the complexity of prescribing opioids. By the spring of 2004, federal agencies including the Food and Drug Administration (FDA), the Drug Enforcement Administration (DEA) and the White House Office of National Drug Policy launched coordinated prescription drug abuse prevention initiatives that emphasized ed-

Barth L. Wilsey, Scott M. Fishman and Christine Ogden

Prescription Opioid Abuse in the Emergency Department

Barth L. Wilsey, Scott M. Fishman and Christine Ogden

Barth L. Wilsey, M.D., is a Joint Appointee of the Sacramento Veterans Administration and the University of California, Davis and is Director of the VANCHCS/UCD Analgesic Research Center. Scott M. Fishman, Ph.D., is Chief of the Division of Pain Medicine and Professor of Anesthesiology and Pain Medicine at the University of California, Davis. Dr. Fishman is also President of the American Academy of Pain Medicine. Christine Ogden, B.S., was previously a research associate in the Department of Anesthesiology and Pain Medicine at the University of California, Davis. She is now a first year medical student at the University of California, Davis.
uccating physicians about how to properly assess and monitor patients on long-term opioids. At the same time, it was acknowledged that patient care should not be undermined as law enforcement officials act to curb illicit activities. One area of concern has been opioid prescribing and drug-seeking in the ED. In order to avoid inappropriate prescribing, ED physicians must be able to confidently assess the potential of opioid misuse. But these physicians often lack the skill or the time for a thorough assessment. There is also a lack of uniformity in the approach to chronic pain patients by ED physicians. In one study, judgments regarding prescribing practices were highly variable; under an identical scenario 10% of physicians were against prescribing opioids and 10% were likely to prescribe opioids. Furthermore, the same clinical information, such as a patient requesting a strong analgesic, changed the willingness to prescribe opioids in opposite directions for different physicians. Obviously, the role of opioid prescribing in the ED should be on a more logical footing.

The difficulties involved in deciphering the reasonableness of a particular patient’s request for opioids in the ED are compounded by the high prevalence of addiction in these patients. Although data is far from conclusive, there is concern that a considerable number of patients presenting to emergency rooms, as much as half in some settings, have a history of substance abuse. As a history of addiction has been associated with prescription opioid abuse, aberrant behaviors among chronic pain patients seeking opioids in the ED should not be an unexpected phenomenon. But there is also the concern that patients with chronic pain may receive inadequate analgesia as a result of prejudicial treatment deriving from guilt by association. Differentiating between true addiction and the patient whose analgesic requirements are not being met (e.g., pseudoaddiction) is usually difficult, particularly in the constrained confines of a single ED encounter. Nonetheless, some patients are denied treatment for pain because of the fear that their complaints are “not real.” ED physicians may be concerned that a patient exhibits drug-seeking behavior whereby amplification or falsification of a somatic complaint is offered by a patient in order to receive a prescription for opioids. Alternatively, an ED physician may encounter a patient who reports having allergies to multiple non-narcotic medications. Request for a specific type of pain medication may also raise concerns about possible addiction. A few patients may engage in malingering or another form of deception for secondary gain, complaining of symptoms that are impossible to prove or refute. As the pain of a toothache, migraine headache, back pain, or an abdominal complaint is not directly measurable; there is always subjectivity in the complaint of chronic pain. Certain scenarios may increase the degree of suspicion; a patient may present with a chronic pain problem that is being treated with opioids by their primary care doctor or specialist who is reported to be unavailable to refill the medication. Patients may present stating they are without the opioids that they take regularly and perhaps are now visiting a relative in a distant city and need a prescription on an emergency basis. Consultation with their primary care physician may reveal a history of self-escalation of med-

Figure 1
ications, calling for early refills secondary to lost or stolen prescriptions, multiple telephone calls, and/or visitations without an appointment to obtain medications. Although some emergency physicians may consider themselves adept in the identification of patients with abusive tendencies, this is not always possible. While the evidence may seem self-evident, proving allegations of misuse is usually impossible in a single encounter. At the same time, it is becoming increasingly difficult to defend the practice of denying opioids because of suspicions that pain is inexplicable or that the patient is a “drug-seeker.”

There is surprisingly little data on the exact prevalence of prescription opioid abuse in patients with chronic pain seeking care in the ED. Some have referred to the “distortion perpetrated by a small number of patients that frequent EDs requesting opioids for illicit purposes.” However, literature on the exact pervasiveness of prescription drug abuse in this setting is not available. This is in direct contrast to alcohol and illicit drug abuse, which has been studied extensively in this population. The lack of data on this topic is likely related to the necessity of deciphering subtle signs of prescription opioid abuse over multiple encounters. Previous efforts to characterize this type of conduct have enumerated several telltale themes (Table 1). Obviously, an average ED physician is not going to be able to make this diagnosis from this inventory as they are not going to witness repeated calls or have the occasion to observe a patient returning to the office because they have run out of medications. But screening tools are under development which may prove to be useful in the ED as a surrogate measure to predict the propensity for prescription opioid abuse in a given patient.

**Screening for Prescription Opioid Abuse**

An early method of categorizing aberrant behaviors in patients on a chronic opioid regimen was developed by Chabal, et al. Working in a pain clinic in a Veterans Administration facility, the authors developed a prescription abuse checklist (Table 2). Patients with chronic pain meeting three out of five of these criteria were arbitrarily classified as opioid abusers. Although notable for its relative simplicity (e.g., the criteria could be applied during normal clinic interactions), this method is not without its drawbacks. Most problematic is the fact that some patients may manifest prescription opioid abuse but not have observable aberrant behaviors. This was discovered in a retrospective review of charts in a pain clinic where 21% of seemingly uncomplicated pain patients taking opioids were shown to have urine toxicology results that were suspicious for abuse by virtue of the fact that illicit drugs were found to be present in the samples tested. It is therefore thought that behavioral examinations must be supplemented with toxicology screening in order to detect the true incidence of prescription drug abuse. The combination of these methodologies is, for this present time, probably the “gold standard” upon which other methods, i.e. screening devices, should be compared. But there are other patients who manifest neither behavioral nor toxicologic abnormalities, yet still abuse medications. Abusive behaviors such as selling diverted medications, and/or crushing oral medications and injecting them, are far less likely to be discernable.

Shortly after Chabal, et al., reported their findings, Compton and her psychiatric colleagues reported upon the results of a pilot assessment tool called the Prescription Drug Use Questionnaire. Acting as psy-
chiatry liaison consultants, the authors screened referrals with suspected abusive behaviors. From this data, the authors developed a list of forty-two items, which seemed to be reasonable correlates of prescription opioid abuse. The authors then administered the assessment as a structured interview over approximately twenty minutes. This instrument evaluates the pain condition, opioid use patterns, social and family factors, family history of pain and substance abuse syndromes, patient history of substance abuse, and psychiatric history. Responses to these items by subjects diagnosed as being addicted by an addiction medicine specialist differed significantly from those of non-addicted patients. Those with a substance use disorder were significantly more likely to save or hoard unused medications, to use analgesics to relieve symptoms other than pain (i.e., insomnia, anxiety, depression), to supplement analgesics with alcohol or other psychoactive drugs, and to report having a practitioner previously limit or terminate care due to concerns about the subject’s analgesic use. Although not meeting statistical significance, addicted patients had a pattern of losing prescriptions/medications, had forged a prescription, or had obtained opioids from street sources. To identify a manageable number of items that best predicted the presence of addictive disease, logistic regression analysis was performed. Positive responses to three items correctly classified 93% of subjects as having an addiction. The three questions are presented in Table 3. These personally sensitive questions have not been utilized widely because it would be unlikely that someone would answer affirmatively in a typical clinical setting. Less offensive queries that correlate with prescription opioid abuse would possibly derive more information.

Recently, there have been attempts to discover methods of eliciting accurate information from respondents relating to opioid abuse potential. Adams and colleagues developed a twenty-six item instrument called the Pain Medication Questionnaire (PMQ). This scale was notable for its readability. It was designed to be filled out by the patient and scored by the clinician to determine the relative risk of future prescription opioid abuse. Butler, et al., have reported on their Screener and Opioid Assessment for Patients with Pain (SOAPP). This instrument is also a paper and pencil questionnaire intended to facilitate treatment planning for chronic pain patients being considered for long-term opioid treatment. It has twenty-four items designed to serve as a test for the potential of a patient to develop abusive behaviors. These latter two studies demonstrate the feasibility of employing a self-report instrument completed by a respondent with items somewhat similar to those applied in the structured interviews of Compton. But unlike the latter, the items

<table>
<thead>
<tr>
<th>Table 1</th>
<th>Prescription Opioid Abuse</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Selling prescription drugs</td>
<td></td>
</tr>
<tr>
<td>2. Forging prescription drugs</td>
<td></td>
</tr>
<tr>
<td>3. Stealing drugs</td>
<td></td>
</tr>
<tr>
<td>4. Injecting oral formulations</td>
<td></td>
</tr>
<tr>
<td>5. Obtaining prescription drugs from nonmedical sources</td>
<td></td>
</tr>
<tr>
<td>6. Concurrently abusing alcohol or other illicit drugs with prescription drugs</td>
<td></td>
</tr>
<tr>
<td>7. Escalating doses on multiple occasions or otherwise failing to comply with the prescribed regimen despite warnings</td>
<td></td>
</tr>
<tr>
<td>8. “Losing” prescribed medication on multiple occasions</td>
<td></td>
</tr>
<tr>
<td>9. Repeatedly seeking prescriptions from other clinicians or from emergency rooms without informing the original prescribing physician</td>
<td></td>
</tr>
<tr>
<td>10. Giving evidence of a deterioration in the ability to function (at work, in the family, or socially) that appears to be related to drug use</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 2</th>
<th>Prescription Opioid Abuse in an Office Practice</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Excessive focus on opiate issues during clinic visits</td>
<td></td>
</tr>
<tr>
<td>2. A pattern of early refills or dose escalation in the absence of clinical change</td>
<td></td>
</tr>
<tr>
<td>3. Multiple phone calls or visits about opiate prescriptions</td>
<td></td>
</tr>
<tr>
<td>4. A pattern of prescription problems</td>
<td></td>
</tr>
<tr>
<td>5. Supplemental sources of opiates</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Table 3</th>
<th>Three Questions from Prescription Drug Use Questionnaire (PDUQ)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Does the patient believe that he/she is addicted to opioid analgesics?</td>
<td></td>
</tr>
<tr>
<td>2. Is there a pattern of the patient increasing prescribed analgesic dose or frequency?</td>
<td></td>
</tr>
<tr>
<td>3. Does the patient have preferences for specific analgesics and/or routes of administration (i.e., IV, IM routes over oral)?</td>
<td></td>
</tr>
</tbody>
</table>
would appear to be less objectionable to answer. In addition to these screening surveys, there is also a new tool for ongoing assessment. Passik, et al., have described their Pain Assessment and Documentation Tool (PADT). This tool captures the relevant outcomes for pain management: analgesia, activities of daily living, adverse events, and aberrant drug-taking behaviors. The mnemonic “4 As” remind clinicians that pain therapy can be considered successful only if it provides pain relief (analgesia) and stabilizes or improves psychosocial functioning (activities of daily living) without compromising important areas of functioning (adverse events) or triggering abuse (aberrant drug-taking behaviors).

The major drawback to the clinical utilization of any of the aforementioned instruments in the ED is the length of administration; each assessment takes several minutes to complete. Undoubtedly, differences in the length of a screening procedure weigh heavily upon the burden of clinicians and patients. Others have suggested abbreviated questionnaires. For instance, officials at the National Institute of Drug Abuse (NIDA) recommended a modification of the CAGE Questionnaire that is commonly used with alcoholism. Adapting this short questionnaire from Ewing, they substituted “prescription drugs” for the word “alcohol” (Table 4). Although a tempting use of a validated screening instrument that is familiar to most ED practitioners, the validity of the revised questionnaire has not been established. Another abbreviated questionnaire was suggested by the work of Michna, et al., who found that chronic pain patients attending a pain clinic had certain characteristic that predicted subsequent aberrant behavior. Patients who admitted to a family history of substance abuse, a history of legal problems, and drug or alcohol abuse were prone to demonstrate more aberrant opioid-related behaviors. Such behaviors include a higher incidence of lost or stolen prescriptions and the presence of illicit substances in their urine. This study demonstrated that questions about abuse history and legal problems can be useful in predicting aberrant drug-related behavior with opioid use in persons with chronic pain of nonmalignant origin. The three questions employed are presented in Table 5. If a subject answered two or more of these questions affirmatively, they were considered to be at high risk for prescription opioid abuse. At present, this is probably the most useful means for an ED physician to screen for prescription opioid abuse because it involves querying about a less offensive issue such as a family member’s problems with addiction. Given human nature, a patient is more likely to admit that someone else is abusing or has abused drugs or alcohol than they might be to admit their own frailties. Indirect and somewhat subtle questioning of this sort is likely to uncover a clue that the patient may have a higher-than-normal potential for prescription opioid abuse. Similarly, the question about having a problem with drugs or alcohol or attendance at Alcoholics Anonymous (AA) or Narcotics Anonymous (NA) meetings is likely to elicit a higher end of the spectrum in those individuals who are at risk.

### Table 4

**CAGE Modified for Prescription Opioid Abuse**

1. Have you ever felt the need to cut down on your use of prescription drugs?  
2. Have you ever felt annoyed by remarks your friends or loved ones made about your use of prescription drugs?  
3. Have you ever felt guilty or remorseful about your use of prescription drugs?  
4. Have you ever used prescription drugs as a way to “get going” or to “calm down”?

### Table 5

**Abuse History and Legal Problems**

1. Is there a history of alcohol or substance abuse in your family, even among your grandparents, aunts, or uncles?  
2. Have you ever had a problem with drugs or alcohol or attended Alcoholics Anonymous (AA) or Narcotics Anonymous (NA) meetings?  
3. Have you ever had any legal problems or been charged with driving while intoxicated (DWI) or driving under the influence (DUI)?

Physicians may employ other methods to suspect drug or alcohol misuse. Findings on physical examination include an assessment for the signs associated with addiction including tremor, the odor of alcohol on the patient’s breath, an enlarged, tender liver, nasal irritation (suggestive of cocaine inhalation), conjunctival irritation (suggestive of exposure to marijuana smoke), labile blood pressure, tachycardia (suggestive of alcohol withdrawal), and/or “aftershave/mouthwash” syndrome (masking the odor of alcohol), and the odor of marijuana on clothing. Obviously, the presence or absence of any one of these findings is not diagnostic for drug dependence. Likewise, no laboratory findings exist that are specific for this diagnosis. A urine toxicology screen...
is the perhaps the best examination to confirm involvement with illicit drugs (amphetamines, barbiturates, benzodiazepines, cocaine, opiates, phencyclidine (PCP), and marijuana). But laboratory procedures vary greatly and false positive urine screens are prevalent. For instance, the pain patient may be positive for opioids as a result of medication that they have received from legitimate sources or negative even though they are taking prescribed dosages due to variable thresholds for certain laboratories. In many states, patients in pain may use marijuana for medicinal purposes, although the formidable debate on medical marijuana is probably now over (for the time being) in light of the 2005 U.S. Supreme Court ruling that now essentially bans its use by patients irrespective of state law. Other false positives include positive testing for opioids due to recent ingestion of poppy seeds. Immunoassays are the most commonly used initial screening method and are either laboratory-based or point-of-collection tests. As in pain clinics, the optimal identification of drug abuse in the ED requires both the use of a history and a drug screen.

Since there is no valid characteristic symptom or sign, lab nor screening test that is even marginally capable of differentiating patients in pain who are using their complaint for inappropriate purposes, ED clinicians may perceive these patients as a group that is met with disbelief. Such patients become lumped together, as cases that ED clinicians may be hesitant to treat. Thus, it remains an imperative to discover a simple means of differentiating “legitimate” pain patients from abusers. In the interim, it is important for the ED physician to maintain a compassionate approach to each of these patients and to offer them a trusting environment unless mitigating circumstances develop (i.e., the patient is discovered to have procured prescription opioids from multiple providers).

**Treating the Patient with Concurrent Addiction and Chronic Pain**

For patients who are perceived as being at high risk of having prescription opioid abuse because of concurrent substance abuse, a variety of strategies may be contemplated. Prior to the institution of chronic pain management with abusable medications, it is advisable to consult with either a pain management or addiction specialist. The rationale behind this recommendation is that such patients have resource utilization needs that are typically higher than usual, exceeding the type and extent available in non-specialty programs. Such patients may need a carefully developed and delivered written opioid agreement so they are made aware of the permissible and unacceptable behaviors while on prescription opioids. Moreover, such patients will require a treatment milieu that supports sobriety and has the ability to monitor for the behavioral dysfunction in all domains of life that is the hallmark of addiction.

Despite the best intentions, sometimes it will be necessary for a ED physician to refuse to provide opioids to a patient who has chronic pain combined with an addiction problem. Due to insurance issues and the long waiting period to be seen by some specialty clinics, it may be the path of least resistance for this patient to fall into a cycle of returning to the ED for their pain medications/addiction, offering no real solution to either the pain problem nor their substance abuse. To avoid this conundrum, referral to the ED Social Services Department would be warranted to see if placement could be arranged to avoid continuing dependence upon the overburdened ED.

Many physicians consider a history of substance abuse to be a relative contraindication to chronic opioid therapy for nonmalignant pain because of a predilection for relapse. Dunbar and Katz examined twenty patients with a history of substance abuse treated with chronic opioid therapy for nonmalignant pain. This retrospective study looked at the predictive factors associated with prescription abuse. The subgroup that abused opioids were much more likely to have had recent histories of polysubstance abuse. They also had a tendency to abuse medications early on in their clinical course; requesting early refills soon after initiating opioids. It is notable that signing an opioid contract was not in and of itself a deterrent for prescription drug abuse in this retrospective study. Patients in the subgroup that did not manifest prescription opioid abuse were more likely to have a history of remote alcohol abuse. Especially noticeable among those who fared well were patients who were active members of AA and/or those with a stable support system (e.g., family).

It is widely held that prescribing chronic dosing of controlled substances to practicing addicts (ongoing active substance abuse) is inappropriate because of a significant risk of enabling the dysfunctional behavior leading to further harm. Actively abusing addicts with
chronic pain are usually best managed as an inpatient in a drug treatment facility where they receive their medication for pain as well as treatment for addiction in a controlled setting. Inasmuch as 30-80% of substance abusers suffer from co-existing psychiatric disorders, psychiatric evaluation and treatment, if indicated, should be implemented at the initiation of this controlled therapy. Following discharge, patients should see the prescribing physician, and if the decision is made to use opioids in the setting of recent or remote substance abuse, additional opioids for pain may be given every few days with the goal of progressively lengthening the interval between visits as mutual trust develops. The physician can develop a trusting relationship by demonstrating empathy while at the same time establishing clear behavioral boundaries. During the opioid trial, the patient must demonstrate compliance with all aspects of the treatment regimen. It is often recommended that prescription bottles be brought to the clinic so that the clinician may perform pill counts or, if transdermal patches are used, they may be returned after use for inspection to determine if they have not been tampered. Periodic urine toxicology screens should be performed to exclude illicit drug use. The patient should be informed at the onset of treatment with opioids that the pain medications will not eliminate their pain entirely. Instead, a treatment plan can be prearranged whereby a percentage of pain relief (e.g., approximately 50%) will be the primary goal which should be manifest through evidence of improved physical and psychological functionality. Functional outcomes are set at the beginning of treatment and usually involve determining the means by which functional gains may be made evident to the clinician. For example, patients may bring in collateral sources such as family members or documentation of participation in physical or psychological reconditioning activities or sobriety support programs. This script is then religiously followed during each subsequent visit. Using a mutually agreed upon course, a successful trial should result in the reduction of the drug-seeking behavior normally encountered in this setting as the patient notes advanced pain relief and improved function. In addition to office visits, enrollment in a support group for substance abusers or other programs that promote improved social, psychological, or physical function are mandatory. The type of organized meeting or program should be left up to the preference of the patient and physician. Some prefer alcoholics anonymous (A.A.) or narcotics anonymous (N.A.) while others utilize group psychotherapy. But the relative importance of this type of support activity is well known and cannot be overemphasized. The same type of close prescribing physician supervision with complementary social, psychological, and physical support can be utilized in the case of a person with a remote history of substance abuse; the only difference being that this latter individual may be manageable on an outpatient basis at the time of initiation of contact.

Influence of Opioid Contracting on Prescription Drug Compliance

Contracts are often employed in the chronic administration of opioids and are intended to improve adherence to a treatment regimen. In addition to enhancement of compliance, contracts provide education and documentation. Although the efficacy of opioid contracting is not known, studies reviewing the use of contracts for patients in methadone programs are promising. Nolimal, et al., studied fourteen methadone maintenance patients who were offered either discharge or a contract. The rate of drug-free urine toxicology screens increased from 38% in the three months before intervention to 55% during the three-month intervention with contracting. Marked improvement occurred during the first month, with the effect waning thereafter. Saxon, et al., evaluated the outcome of a mandatory structured contingency contracting system in a methadone maintenance program in a group of patients who continued illicit drug use. The contracting involved weekly urine toxicology screens. If patients continued the use of illicit drugs, the initial contingency was to lower their methadone dose. Tapering (detoxification) and discharge followed subsequent violations. Illicit opioid use decreased significantly for subjects utilizing the more stringent contracts. There is limited evidence to support successful contracting in the pain population. An anecdotal report by Burchman, et al., described the implementation of a successful agreement in patients with non-malignant pain. Key features included acknowledgement that previous treatment strategies had failed, a listing of side effects and the risks of opioid therapy (including the potential for addiction), and the contingencies of treatment, including the importance of pain relief coupled to enhanced function via active participation in other therapies. Fishman, et al., compared thirty-nine opioid contracts from major academic programs finding wide variability of content. However, there was a core group of themes found consistently amongst the contracts reviewed. The “opioid contract” often included a clear description of what constitutes medication use and abuse, terms for random drug screening, consequences of contract violations, and measures for opioid discontinuation should this become required. Multiple instances of minor deviations of the contract were generally tolerated before resorting to severing the agreement. However, unlawful activities such as
forging prescriptions, selling drugs and/or resumption of alcohol or illicit drug intake or abuse were often grounds for immediate tapering and discharge.

A contract is often used to supplement the introduction of important concepts required for safe opioid prescribing to patients beyond that which the busy office physician can accomplish during a visit. It may behoove the ED physician to inquire about the presence of an opioid agreement which may provide information as to whether or not the patient had been under the care of a physician who tried to prevent the patient from seeking multiple providers. It might also be a means of deciphering whether or not the patient has had difficulties with other providers who might have stopped providing opioids. But there may have been other reasons for a treatment failure. Previous aborted therapy with opioids may have had any number of underlying reasons, including: (1) inadequate titration of dose, (2) unsuitable dosing schedule, (3) opioid insensitive pain relating to the nature of the pain generator (e.g., neuropathic pain), and (4) development of side-effects that limit dose escalation. Therefore, the practitioner cannot assume a cause and effect relationship from recognition of aberrant behaviors to treatment failure.

Influence of the Type of Medication on Prescription Drug Compliance

A number of opioids are available for clinical use, including both short and long-acting agents (Table 6). The types of opioids and vehicles for drug administration are expanding as basic scientists and pharmaceutical companies recognize the need for alternative agents and different methods of drug delivery. New and more potent synthetic opioids with novel delivery systems have been developed and are undergoing clinical trials. A matchstick-sized implanted osmotic pump (Chronogesic™) that delivers sufentanil subcutaneously for more than ninety days is being developed to treat chronic pain.32 Respimat™ is a novel, compact, propellant-free, multi-dose inhaler that employs a spring to push drug solution through a nozzle, which generates a slow-moving aerosol allowing rapid delivery of opioid.33 E-TRANS™ fentanyl, an electrotransport system (ETS), allows self-titration of the drug by the patient, for the management of acute pain with a titratable dose.34 A novel, once-daily, osmotic technology (OROS™), allows extended-release hydromorphone delivery.35 Having a variety of agents available allows opioid rotation; a tactic often employed to (1) rotate between different long-acting opioids to improve analgesia and reduce side-effects, and (2) rotate from a short-acting opioid to a long-acting opioid to establish stable analgesia in order to minimize withdrawal symptoms, risk of tolerance, and addiction. When studied empirically, opioid rotation between different long-acting opioids resulted in better overall analgesia and fewer side-effects at dose levels predicted to be equianalgesic.36 Interestingly, the same study revealed that the majority of the patients rotated from short-acting opioids to long-acting opioids also obtained improved analgesia, but at a cost of a 74% increase in the opioid dose.

The Oregon Health Resources Commission performed an evidenced-based review comparing long and

<table>
<thead>
<tr>
<th>Opioid</th>
<th>Generic Name</th>
<th>Equianalgesic Amount</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Short-Acting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Codeine</td>
<td>200 mg</td>
<td>Most widely employed naturally occurring opioid</td>
<td></td>
</tr>
<tr>
<td>Hydrocodone</td>
<td>20-30 mg</td>
<td>Many products combining hydrocodone and non-opioid analgesics available</td>
<td></td>
</tr>
<tr>
<td>Oxycodone</td>
<td>20-30 mg</td>
<td>High abuse potential</td>
<td></td>
</tr>
<tr>
<td>Propoxyphene</td>
<td>130 mg</td>
<td>Not more effective than aspirin alone</td>
<td></td>
</tr>
<tr>
<td>Tramadol</td>
<td>120 mg</td>
<td>Avoid in patients at risk for seizures or taking SSRIs</td>
<td></td>
</tr>
<tr>
<td>Long-Acting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fentanyl Patch</td>
<td>N/A</td>
<td>25 ug patch is equal to approximately 45 mg of morphine equivalents per day</td>
<td></td>
</tr>
<tr>
<td>Slow Release Morphine</td>
<td>30 mg</td>
<td>Available as 15, 30, 60, and 100 mg pills</td>
<td></td>
</tr>
<tr>
<td>Slow Release Oxycodone</td>
<td>20 mg</td>
<td>Available as 20, 40, and 80 mg pills</td>
<td></td>
</tr>
<tr>
<td>Methadone</td>
<td>10 mg</td>
<td>The morphine-to-methadone conversion ratio increases as the previous dose of morphine increases; conversion ratios may not be bidirectional</td>
<td></td>
</tr>
</tbody>
</table>

Table 6
Equianalgesic Dosing (Morphine 10 mg IV or 30 mg po)
short-acting opioids. No significant differences in efficacy (pain relief) were found and there was insufficient evidence to suggest that long-acting opioids, as a class, were more effective or safer than short-acting opioids in terms of abuse liability. This conclusion is in contradistinction to empirical evidence from a Substance Abuse and Mental Health Services Administration (SAMSHA) Survey (Fig 3) in 2002 revealing that short-acting opioids have a much higher incidence of diversion and therefore, are preferred as agents of abuse. It would seem plausible that chronic pain patients are susceptible to similar inducements for abusive behavior as were the nonmedical users of prescription-type drugs depicted in Fig 3. Therefore, it has been argued on theoretical grounds that long-acting opioids are preferable to the short-acting variety when treating chronic pain. ED physicians may elect to provide either type of medication. It is not critical that the long-acting agents be given at every encounter as the urgent and brief visits that take place in the ED may suffice with the provision of short-acting opioids for brief periods of time. Presumably, the amount of opioid provided should be limited to that which would provide a bridge to the patient’s next appointment with their treating physician who will provide continuity of care. One exception to this general rule would be the patient who manifests an active addictive state. Unless significant mitigating circumstances are present, (e.g., advised by primary care physician to proceed with opioid refill and patient has evidence of strong social support), this patient should not be given opioids by the ED physician. Instead, if possible, the patient should be referred to an addictionist and the patient provided with an opioid taper plus adjuvant analgesics (acetaminophen, NSAIDs, etc.).

**Medical Records Documentation**

Documentation is an essential component of safe and effective opioid prescribing. This point is reflected in the *Model Policy for the Use of Controlled Substances for the Treatment of Pain* by the Federation of State Medical Boards of the United States (http://www.fsmb.org/). The Model Policy is designed to communicate critical messages to licensees including:

- The state medical board views pain management to be important and integral to the practice of medicine
- Opioid analgesics may be necessary for the relief of pain
- The use of opioids for other than legitimate medical purposes poses a threat to the individual and society
- Physicians have a responsibility to minimize the potential for the abuse and diversion of controlled substances
- Physicians will not be sanctioned solely for prescribing opioid analgesics for legitimate medical purposes

This and other guidelines and polices for the use of opioids in the treatment of pain uniformly stress the im-
portance that records remain current, maintained in an accessible manner, and readily available for review. Furthermore, accurate and complete records must include details of the medical history and physical examination, diagnostic, therapeutic and laboratory results, evaluations and consultations, treatment objectives, discussion of risks and benefits, informed consent, treatments, medications (including date, type, dosage, and quantity prescribed), instructions, agreements, and periodic reviews. Although these guidelines are identical to what is expected of any medication treatment, they have not been followed. In an evaluation of 300 randomly selected charts\(^9\) from a Veterans Administration primary care facility, only 39% of the cases were found to have comments pertaining to a treatment plan or the efficacy of prescribed opioids in reducing pain, improving function, or otherwise benefiting the patients. Only 41% of these patients had a physical exam which seemed to be directed at the painful area within the preceding six months. In addition, a paltry 17% of patients receiving opioids had both comments on a treatment plan or follow-up and a pain-related physical exam within six months of chart review. Surprisingly, many patients had repeated documentation of lung and heart auscultation despite a lack of pulmonary or cardiac symptoms but did not have a documented physical exam directed at evaluating the pain complaint. The presence or absence of side-effects was documented in only 9% of the charts reviewed. Another evaluation\(^40\) was conducted at an academic family medicine center in Oregon, examining the percentage of records exhibiting documentation compliance with state prescribing laws and other features indicative of a high standard of care. The Oregon legislature had previously mandated that patients receive a full disclosure, identified as a “Required Material Risk Notification Form,” explaining the risks and benefits of opioid prescribing.\(^41\) Despite having this state law that requiring these consent documents, medications (including date, type, dosage, and quantity prescribed), instructions, agreements, and periodic reviews. Although these guidelines are identical to what is expected of any medication treatment, they have not been followed. In an evaluation of 300 randomly selected charts\(^9\) from a Veterans Administration primary care facility, only 39% of the cases were found to have comments pertaining to a treatment plan or the efficacy of prescribed opioids in reducing pain, improving function, or otherwise benefiting the patients. Only 41% of these patients had a physical exam which seemed to be directed at the painful area within the preceding six months. In addition, a paltry 17% of patients receiving opioids had both comments on a treatment plan or follow-up and a pain-related physical exam within six months of chart review. Surprisingly, many patients had repeated documentation of lung and heart auscultation despite a lack of pulmonary or cardiac symptoms but did not have a documented physical exam directed at evaluating the pain complaint. The presence or absence of side-effects was documented in only 9% of the charts reviewed. Another evaluation\(^40\) was conducted at an academic family medicine center in Oregon, examining the percentage of records exhibiting documentation compliance with state prescribing laws and other features indicative of a high standard of care. The Oregon legislature had previously mandated that patients receive a full disclosure, identified as a “Required Material Risk Notification Form,” explaining the risks and benefits of opioid prescribing.\(^41\) Despite having this state law that requiring these consent documents, they were absent from 100% of charts reviewed by the authors. Similar to the findings at the Veterans Administration facility,\(^42\) medication contracts were only present in 39% of records and documentation of a pain evaluation and functional evaluation was present in 67% and 54% of records, respectively.

Obviously, the ED physician cannot be expected to provide the degree of documentation that a continuity care physician can accumulate during multiple encounters. Nonetheless, the propagation and promotion of guidelines by various professional and governmental agencies have generally converged on the importance of documentation in opioid therapy. Fears of punitive consequences for not meeting such demanding requirements, coupled with low yield in return for time and effort invested in conforming to these guidelines, have probably deterred physicians from prescribing opioids even in appropriate settings. The results of these fears are continued pervasive inadequacies of pain management. This “opiophobia” promotes a cycle that perpetuates barriers to care throughout the medical sphere including primary care, specialty, and emergency settings. In order to circumvent this in the ED setting, medical record documentation of opioid prescribing should be a simple explanation of the rationale for the prescription with a disclosure of the amount and type of medication.

### The Role of Habitual Patient Files and Prescription Monitoring Programs in Assessing Prescription Opioid Abuse

EDs may maintain an inventory of names of patients who are suspected of repeatedly visiting to obtain prescription opioids for nontherapeutic purposes. One purported benefit of “habitual patient files,” “frequent flyer files,” and “special needs files” is the potential for reducing prescription fraud. It is argued that habitual patient files might support physicians in complying with the prohibition on prescribing to addicts.\(^43\) Maintenance of such files present many possible benefits, but they also may stigmatize chronic pain patients in the ED. Such patients may wonder why they have permanent files while others do not. Such files may also be used to negatively label these patents as different and potentially decrease access to care. While in principle this may be considered akin to the office-based physician’s recollection or notation of aberrant behaviors in the medical record, it is different because of issues concerning patient confidentiality. New federal privacy regulations implemented under the Health Insurance Portability and Accountability Act of 1996 (HIPAA) reflect the concern that newer modes of electronic transmission of health (and insurance) data may be susceptible to disclosure through security breaches. When considering instituting habitual patient files, it is suggested that a health-care attorney with expertise in privacy issues be consulted to assure that the process conforms to both state and federal law. In general, habitual patient files are permissible if their goals include “protecting patients from harm as the result of drug abuse, preventing the inappropriate use of valuable ED resources, or protecting society from harms caused by the resale of ill-gotten drugs or the actions of intoxicated persons.”\(^44\) HIPAA allows physicians to share protected health information with other physicians, nurses, and other health-care workers for the purposes of treatment. But the same information cannot be provided outside of the covered entity (i.e., hospital, healthcare plan, etc.) so that the contents of a habitual file should...
never be divulged to health-care workers in another institution. The design and implementation of the file should optimally have specific provisions for managing chronic pain patients. For instance, there might be limits on the number of visits or the amount of medication to be provided to chronic pain patients. This would be systematized as practice parameters by the ED medical staff to be repeatedly amended through continuous quality improvement so that newer information regarding opioid prescribing may be introduced. The opportunity also exists for EDs to provide a priori guidance as to the type of pain medications that their institution will provide. In this manner, the use of methadone might be limited. Although methadone has numerous advantages as an analgesic, its unusual metabolism distinguishes it as a potentially unsafe medication. Any physician prescribing this medication for chronic pain has to be aware of its interactions with other medications as well as the tendency for this drug to accumulate and cause sedation and respiratory depression even in individuals who are tolerant to other opioids. The ED medical staff might also consider listing providers that will accept chronic pain patients into their practices so that referrals to appropriate resources may be enhanced.

There have been legislative enactments to allow disclosure of information between covered entities to prevent the practice of doctor shopping, where abusers visit multiple physicians in order to obtain several prescriptions. The Kentucky All Schedules Prescription Electronic Reporting (KASPER) system archives information by having pharmacists enter the controlled substance prescription information into a database. Practitioners are then provided real time access to this information. However, doctor shoppers have circumvented KASPER by traveling to one of the seven states surrounding Kentucky. As a partial attempt to remedy this problem, The National All Schedules Prescription Electronic Reporting (NASPER) bill has been passed into law by the U.S. Congress and is awaiting the President’s signature. This bill offers a voluntary program whereby independent states may receive federal funding for a prescription monitoring program. NASPER does not allow for a national database on the dispensing of controlled substances in Schedules II-IV and thus falls short of the potential for disclosure of doctor shopping outside of the boundaries of a single state. But it is an initial attempt by the federal government to manage doctor shopping. Whether or not an individual ED physician would have access to this data will depend upon the state in which he or she practiced since the present plan is to provide money to individual states allowing each to formulate their own individualized program.

Conclusions

Patients with chronic pain may present to the ED in hopes of receiving treatment. Concerns range from whether some patient populations receive inadequate analgesia to whether some patients receive too much opioid medication, potentially placing providers under scrutiny. Our empirical evaluation of the scope of this problem (in press) revealed that many of the patients who utilized either a metropolitan University or a Veterans Administration ED had a high propensity for prescription opioid abuse. This was corroborated by a similarly high incidence of a history of current or past substance abuse in these patients. Multivariate analysis revealed a positive correlation of the prescription opioid abuse with several psychological factors (e.g., anxiety and personality disorders) that are commonly associated with chronic pain.

At the present time, there is no simple test to determine if a patient’s pain complaint is legitimate or if there are additional risks such as whether or not the medications have been prescribed by multiple providers. Physicians should evaluate chronic pain patients for a personal or family history of alcoholism, substance abuse, and legal problems (i.e., driving under the influence). If the majority of these items are answered affirmatively, there is a high likelihood that the patient will manifest some type of aberrant behavior. This does not mean that the patient should not be provided pain medications. But a notation can be placed in the medical record that the patient is at high risk of prescription opioid abuse. If a patient’s history of alcoholism or substance abuse is current, a referral to an addictionist or pain practitioner should be made and it may be wise to avoid providing opioids in the ED. As alternatives, NSAIDs and other medicines that have been found to be effective substitutes for opioids in treating headache, sickle cell crisis, and renal colic may be prescribed. If the problem with abuse is remote, then the ED physician must use his/her best judgment as to whether or not the patient can be provided opioids. If there is no history of abuse, then the ED physician can conclude that the patient is at low risk for manifesting abusive behaviors. In either case, the patient should be counseled to seek the care of a practitioner who can provide continuity of care. Habitual files can be used to provide a means for an ED to structure treatment of chronic pain patients and maintain continuity of information. In the future, Prescription Monitoring Programs may replace habitual files so that information of the use of multiple providers is available on a real time basis. Until then, the prudent ED physician must balance compassion with attempts to avoid deception.
Acknowledgements
The authors gratefully acknowledge the support of the Mayday Fund in the development of this article.

References
20. Supra note 7.
23. United States Supreme Court decision in Gonzales v. Raich 545 U.S. 267, 125 S. Ct. 2195 (2005).


42. Supra note 39.


50. P. Compton, etc., supra note 13.

51. NIDA, supra note 18.

52. M. C. Reid, etc., supra note 7.
The prevalence of pain in patients presenting to Emergency Departments (ED) has been well documented by both Cordell and Johnston. Equally well documented has been the apparent failure to adequately control that pain. In 1990 Selbst found that patients with long bone fractures received little analgesia in the ED, and Ngai et al., showed that the under-treatment of pain continued after discharge. In a prospective study, Ducharme and Barber found that up to one third of patients presented with severe pain and were often unrelieved at discharge. Even though specific patient subgroups appear to be at greater risk, all patients are potential victims of oligoanalgesia – the under-treatment of pain. Despite an ever increasing volume of research about pain in emergency medicine, dissemination of relevant information with widespread change in practice patterns has not been witnessed. Recent studies continue to affirm that pain management in the ED is suboptimal.

Unlike the development of a clinical pathway (see Table 1 for definitions) or a protocol, both of which focus on specific care for a particular medical condition, altering attitudes toward pain management is much more difficult. While it can be agreed in general that severe pain should be controlled, it is difficult, if not impossible, to establish beforehand what degree of control is optimal for any one patient. Patient selected endpoints such as amount of pain relief or final pain score vary widely from individual to individual. Furthermore, physician (and nurse) disbelief in patient reporting of pain, opiophobia and fear of drug seeking are persistent subjective barriers resistant to objective recommendations.

In an effort to achieve a “pain free” ED, one of the options is to modify clinical practice with the implementation of clinical practice guidelines. Guidelines are best accepted when based on meta-analyses, or other high-quality data. Such research is still lacking in many areas of pain management and will continue to be lacking for the foreseeable future. Implementation of guidelines has been met with considerable resistance by physicians. There are concerns that guidelines may be established as a standard of care, placing physicians at legal risk if guidelines are not rigidly followed. These concerns arise because many clinical guidelines lack objective graded evidence, relying on “expert opinion.” Frequently, participants in guideline creation are re-

---

**Clinical Guidelines and Policies: Can they Improve Emergency Department Pain Management?**

*James Ducharme*

---

*James Ducharme, M.D., is a Professor in Emergency Medicine at Dalhousie University in Canada. He is also the Clinical Director of the Department of Emergency Medicine at the Atlantic Health Sciences Corporation in Saint John New Brunswick Canada. He is co-editor of the recently published book* *Pain Management and Sedation: Emergency Department Management.*
ceiving pharmaceutical funding, introducing a potential bias that may influence opinion and prevent objective interpretation of available data.\textsuperscript{10} Further resistance to guideline implementation may arise when national associations produce guidelines that conflict with each other. As an example, in one guideline discussing fever without a source in children, the American Academy of Pediatrics (AAP) and the American College of Physicians (ACEP) state that otitis media should not be considered a source of that fever.\textsuperscript{11} Paradoxically, in its most recent guideline on otitis media, the AAP states that treatment for otitis media may be considered optional unless severe, as suggested by the presence of a fever greater than 38.5 degrees Celsius.\textsuperscript{12} Distinct organizations may produce conflicting recommendations for the same topic, leaving the clinician uncertain which to follow.

In the following pages, an effort will be made to first review the guidelines and recommendations published by national emergency medicine associations that are related to pain management. Barriers to guideline implementation will be reviewed, hopefully allowing identification of ways to alter care. There will be discussion about the legal concerns toward guidelines that prevent their implementation, followed by the risks of relying on meta-analyses to develop such guidelines. Improving the (poor) dissemination of new medical information will also be discussed briefly. After identifying these barriers, methods that have succeeded in improving pain management in other areas of health care will be reviewed in an effort to identify how guidelines or pathways can be developed that will improve pain management in emergency medicine.

**Historical Overview**

In an effort to improve the approach to pain management, The Canadian Association of Emergency Physicians (CAEP) produced a consensus document in 1994.\textsuperscript{13} No formal guideline could be developed due to inadequate data and lack of randomized trials comparing treatment options. Instead, suggestions derived by a consensus group from available literature were proposed. These suggestions discussed both pain evaluation and management. A decade later, data still does not allow the production of an evidence-based guideline that could define a general approach to emergency patients in pain. Rather, more focused guidelines dealing with individual topics are required.

Three pain-related guidelines have been produced by ACEP. ACEP first produced a clinical practice guideline about chest pain in 1995.\textsuperscript{14} This guideline attempted to present an algorithmic approach that evaluated and treated patients for potential ischemic heart disease. The rigidity and complexity of the guideline seemed to be deterrents: a year after publication almost half of emergency physicians were aware of its existence, yet only 12% felt it had modified their clinical practice.\textsuperscript{15} ACEP has also produced guidelines with respect to headache\textsuperscript{16} and non-traumatic abdominal pain.\textsuperscript{17} Within each of these guidelines specific questions have been answered, each answer based on graded evidence. The recommendations produced concentrated on the investigation and treatment of various illnesses causing headache, while in the policy discussing abdominal pain, only one paragraph reviewed the use of opioids for pain control. No specific recommendations for pain management have been made in any of the ACEP policies. In the headache policy, physicians were cautioned to not rely on response to medication as a diagnostic indicator.

The 1999 CAEP guideline for migraine management provided graded, evidence-based recommendations with respect to abortive medications.\textsuperscript{18} No recommendations for optimal treatment were (and still are not) possible due to an absence of comparative studies. In the same year CAEP also published the Canadian Emergency Department Triage and Acuity Scale
Pain evaluation, using a numerical rating scale, was integrated into CTAS. All patients with severe (a score of greater than seven, using a scale from zero to ten) visceral chest pain, abdominal pain or headache were expected to be seen by a physician within fifteen minutes of registering, while all other patients with severe pain had to be seen within thirty minutes. Unfortunately many nurses do not complete the pain scale, with less than 50% compliance in the author’s department. Overcrowding further prevents emergency departments from meeting the time expectations established by these recommendations. Despite accreditation for Canadian emergency departments partially dependent on fractile response times established by CTAS, it is the author’s experience almost no major ED has been able to successfully meet these pain-related expectations.

Similarly, the Joint Commission on the Accreditation of Healthcare Organizations (JCAHO) mandated pain evaluation at both registration and discharge.20

A physician should neither be deemed negligent simply for departing from a guideline’s recommendations nor be deemed innocent for adhering to them.

This recommendation was not evidence-based: there is inadequate evidence to say if such documentation improves pain management.21 It is also uncertain that such a “top-down” approach can succeed in changing the health care beliefs that are preventing better pain control at present.22

Dr. Anne-Maree Kelly has attempted to initiate a nurse-driven approach to pain management in Australia.23 Such an approach has not only shortened time to pain relief, but also converted the route of delivery from an intramuscular to an intravenous one. Although a nation-wide recommendation may not succeed, it is perhaps possible to successfully overcome recognized pain barriers with local initiatives such as these. Involving healthcare workers in developing their own policy on pain management may ensure buy-in.

In contrast, both ACEP24 and CAEP25 have produced practical clinical guidelines about procedural sedation. Recommendations about monitoring, patients at risk, and discharge criteria have been defined. While both caution against over-sedation, neither provide instructions with respect to medication dosing or endpoints. The apparent focus of both associations was patient safety and medicolegal concerns, rather than defining optimal medication usage for patient comfort. In the 2005 edition of its sedation policy, ACEP invited emergency nurses to collaborate with physicians in the working group that drafted the document, while CAEP attempted to collaborate with the Canadian Anesthesiology Society in the preparation of theirs. The latter was willing to aid in identifying “problems, perspectives and controversial issues,” but was unwilling to co-author the final guideline.26

Guidelines Viewed as Standard of Care: Legal Concerns

Clinical practice guidelines are meant to aid in decision-making with respect to patient evaluation and care. Regardless of the final format, they are aimed at promoting both higher quality and more cost-effective health care by making the clinical knowledge generated through outcomes research available and more easily accessible.27 Well written guidelines should provide some protection from liability: “in a just and rational world, physicians who follow peer reviewed professional guidelines will be held blameless when a patient has bad luck.”28 Our (medical) world is neither just nor rational, unfortunately; physicians remain afraid of being sued.

Increasing numbers of guidelines – with the largest repository at the National Guideline Clearinghouse (NGC) (www.guideline.gov/index.asp) – are slowly impacting the definition of the standard of care. Even so, courts will most likely treat practice guidelines as one piece of evidence in establishing the standard of care, rather than as a primary determinant.29 In their ideal form, guidelines provide a framework for making decisions. Physicians have legitimate concerns that their discretion and judgment – essential in diagnosing and treating individuals – will be eliminated by rigid adherence to guidelines. A physician should neither be deemed negligent simply for departing from a guideline’s recommendations nor be deemed innocent for adhering to them. Since it is questionable at best that a guideline could be considered standard of care for one particular patient, “not only is it not desirable for guidelines to be deemed conclusive, it is not remotely possible in practical terms.”30

Making use of guidelines to establish standard of care becomes more difficult when one recognizes the complexity of the situation. When using the search terms “acute pain” and “emergency medicine” at the NGC, 92 guidelines are found. Some have differing conclusions with respect to the same topic, while many rely on expert opinion due to lack of high quality scientific evidence. There are almost no systematic reviews related to pain evaluation and management. Even when evidence is available, interpretation of the data may vary, leading to discordant recommendations. Expert opin-
ion can be influenced in many ways: assumptions may be outdated or personal biases from clinical experience, self-interest or fatigue may be present. It seems therefore that “the veracity of practice guidelines cannot be accepted on face value.”

Before establishing guidelines as a standard of care, other legal issues must also be considered. If there is more than one recognized course of treatment, most courts will allow some flexibility in what is regarded as customary. Standard of care becomes more uncertain when competing guidelines produced by different associations do not make similar recommendations. Conflicting recommendations may occur when different medical bodies disagree, but can also occur when medical recommendations conflict with healthcare organization guidelines more concerned with cost-saving approaches. This may also appear when policies consider societal needs ahead of personal ones. Another legal question may be how courts would rule if a guideline is modified by a particular institution. A smaller medical center, for example, may not be able to apply recommendations derived from tertiary-care academic studies without such local modification. Paradoxically, updating guidelines with new information is more feasible at a local level: many national association guidelines may not be renewed for years. The legal system should thus encourage the use of practice guidelines without discouraging individual judgment or local modification and interpretation. There needs to be a middle ground where physician judgment and the patient’s presenting symptoms can determine the treatment plan. Many times a guideline will be “too blunt an instrument...to establish a clear standard of care.”

Health Policy Development

When any specific health issue affects a substantial portion of a community, or involves the use of substantial common resources, it is a public health problem. This can certainly be stated for pain in general and even more for pain seen in the ED. Up to 12% of patients in Canada have suffered high intensity back pain, while greater than 70% of patients registering in the ED complain of pain as a primary symptom. Policy development becomes important for such widespread issues, especially when there is limited evidence of effectiveness, substantial evidence of adverse effects of treatment or poorly disseminated knowledge resulting in less than optimal care. In such situations, many patients may inappropriately receive therapy that is ineffective or deleterious, or both. In emergency medicine, information about pain is lacking in many areas. There is little formal research available on multi-dimensional assessment of pain in the acute setting. There is inadequate data supporting pain scoring as a clinical tool. Optimal analgesic usage derived from comparative studies does not exist in many areas. There is no literature describing the transformation of acute into chronic pain; and we do not know if aggressive pain management can prevent this transformation or only assures greater adverse effects. There are thus many areas of emergency pain management still lacking important information.

Policy development may direct research efforts or allocation of funding for system development to correct such inadequacies. It may also allow for statements that clarify which treatments may be ineffective and should not continue to be supported with public funding. Such policy should originate from high level research, although some remain convinced that policy making should continue to be predominantly based on experience rather than research-based knowledge.

Some of the reasons for not using evidence include:

1. available research evidence may be considered irrelevant. If one looks at Cochrane type meta-analyses, only randomized controlled trials are included as valid evidence. In developing such meta-analyses, all original data from available randomized trials – both published and unpublished – are collated. This data is then graded for quality and homogeneity, to ensure validity of combining data into one analysis about a specific topic. Other study models, such as retrospective or observational studies, are considered to be of too low a level of evidence to provide relevant or adequate data.
2. there may not be agreement of the interpretation of available research.
3. the social environment may not allow acceptance of available research information.

Transforming evidence into policy may be a “demanding task requiring intellectual rigor, discipline, creativity, clinical judgment and skill, organizational savvy and endurance.” Intellectual rigor is often lacking, however, with many of those involved in policy development either unaware of all possible data or lacking in skills required to interpret data that is available. In one study 61% indicated that the executive summary was “the most important component of the systematic review, followed by the conclusion.” Participants were unable to verify the conclusions of the systematic review, be certain that all pertinent studies were included, or ensure that the authors of the review correctly analyzed and interpreted the data from the original studies. The authors concluded that “there was a strong association between use of the reviews and the perception...
that systematic reviews could overcome the barrier of limited critical appraisal skills. It would seem essential to include people with such skills in the development of policy if we wish to succeed in incorporating the best evidence possible. Due to the gaps in knowledge, as well as the existing pain barriers, policy development at this point would probably be best oriented toward development of pain curricula in medical schools, and identifying and funding areas of pain research. Further policy could be directed toward improving the dissemination of new knowledge to the clinician.

Clinical Practice Guidelines: Incorporation into Professional Practice

Guidelines have been unable to modify and standardize practice to the extent expected. The problem appears to lie not with deficiencies of the guidelines themselves, but with the dissemination and acceptance of the information. “Modification of physician behavior has proven to be so daunting an undertaking that no durable resolution of this problem seems likely in the near-term.” There is an ever growing field of research studying interventions to improve professional practice; most interventions used have been shown convincingly to have little or no impact on (medical) behavior modification.

Continuing medical education strategies such as conferences have little direct impact on modifying professional practice, despite being the most popular learning method cited by practitioners. Table 2 lists the various interventions aimed at modifying clinical practice and their degree of effectiveness. It appears that guideline development and passive dissemination – via mailings, in publications or postings on websites – of the guideline is particularly ineffective. Many national associations have traditionally relied on such a passive approach to circulate new guidelines, with resultant poor awareness and utilization.

If we are to incorporate practice guidelines it appears that outreach visits are the most effective, albeit costly, method. The pharmaceutical industry uses this approach with great effectiveness. Trained people meet with physicians in their practice environment to discuss application of new information into their practice. They may also provide feedback on cases seen by the physician. Hospital working groups may incorporate guidelines into clinical pathways as an effective way of modifying clinical practice, but this is time-demanding and requires frequent reviews. Benchmarking of utilization and auditing of changes in patient outcomes is required to ensure local validity of such modifications, but many physicians are unwilling or unable to provide such a commitment, and are wary of others performing such auditing on their behalf.

Why are physicians so reluctant to incorporate new guidelines into their practice? Reasons are many, but include:

1. Complex packaging; guidelines that are lengthy or complex are almost certain to fail.
2. Social and medicolegal context: most Canadian emergency physicians are willing to embrace new guidelines whereas most American physicians are not.
3. Many guidelines rely on expert opinion due to lack of high quality evidence. Such opinion is felt to be biased, often by pharmaceutical ties.
4. Disagreement with guideline recommendations due to difference of interpretation of the available literature.
5. Physicians feel that a guideline either removes independent decision making, or fails to account for patient variance – the “art” of medicine.
6. Many published guidelines conflict in their recommendations, often because the originating organizations have different objectives in mind (e.g. optimal results versus least expensive).
7. Concern that guidelines may be established as a standard of care, placing physicians at legal risk if guidelines are not rigidly followed.

Unless these barriers are addressed, physicians will continue to avoid guideline usage. It may well be that for most clinicians, behavior modification may be impossible. Solutions may only be possible if started early in training, with medical school curricula addressing such attitudinal changes. Such changes could be aimed not just with respect to guidelines in general: attitude and knowledge toward pain management could specifically be targeted.

Table 2

Interventions to modify clinical practice among health professionals

<table>
<thead>
<tr>
<th>Consistently effective</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• educational outreach visits</td>
<td></td>
</tr>
<tr>
<td>• computerized reminders</td>
<td></td>
</tr>
<tr>
<td>• interactive meetings</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Variably effective</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• audit and feedback</td>
<td></td>
</tr>
<tr>
<td>• presentations by local opinion leaders</td>
<td></td>
</tr>
<tr>
<td>• local consensus processes (such as clinical pathway development)</td>
<td></td>
</tr>
<tr>
<td>• patient mediated interventions</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Little or no effect</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>• educational materials, including clinical guidelines, electronic publications</td>
<td></td>
</tr>
<tr>
<td>• didactic educational sessions</td>
<td></td>
</tr>
<tr>
<td>• clinical guidelines and policies</td>
<td></td>
</tr>
</tbody>
</table>
Gallagher suggests that the use of computerized decision support systems may possibly allow incorporation of guidelines into the practice of present clinicians.45 The system should enable clinicians to apply current evidence at the bedside. Those in existence deal most frequently with medication administration. They allow for cost effective care with minimal drug interactions while reducing medication errors and waste. Physician behavior is not permanently modified, for if such systems are removed, physicians rapidly revert to previous patterns. Imposition of a system does not assure education or understanding. In addition, they have not yet been shown to alter patient outcome, while increasing health care costs.46 Even with such reservations and lack of validation, support systems are being widely implemented.

Successful Interventions that have Modified Pain Management
Despite the concerns expressed above, policies have succeeded in improving our approach to pain. The Agency for Health Care and Policy Research (AHCPR) – now known as the Agency for Healthcare Research and Quality – was very successful in the 1990’s at formalizing approaches to patient care. In particular, its publication on low back pain allowed standardization of research and identification of interventions that did or did not work.47 The AHCPR encouraged the establishment of Patient Outcome Research Teams, so that patient oriented results came to the fore.48

The JCAHO recommendations for pain evaluation, while not evidence-based, have forced all emergency departments to at least consider the possibility of pain in all patients.49 Research in the area of pain in emergency medicine is growing each year. Physicians have started to learn and accept that symptom relief does not preclude diagnosis, and is often as important. Donovan, et al., demonstrated many years ago that patient satisfaction was inexplicably high despite poor pain control, primarily because patients expected to suffer.50 The shift of patients to alternative care practitioners – felt to be more empathetic – has transpired due to the perceived insensitivity of physicians toward symptom relief.51 More recently, patients have felt more empowered about their health care, so that poor pain management may translate into lower levels of patient satisfaction. With health care administration increasingly focused on patient satisfaction, system policy will aim to improve any patient care deficiencies that result in sub-optimal satisfaction scoring. We can only hope that inadequate pain relief will be identified as one of those deficiencies.

JCAHO has also mandated standardized monitoring for procedural sedation, irrespective of where such sedation occurs in a hospital. In order to continue to perform procedural sedation, emergency physicians have established national guidelines. The old days of strapping down a child or telling a patient to “tough it out” have thankfully passed, due to such guidelines. It is now considered an integral part of emergency medicine practice to perform procedures safely and painlessly. It would seem that despite concerns that administration dictating care to physicians makes medical practice too rigid, in some instances patient pain management has improved only through such ventures. Further evidence-based guidelines seem justified from the above success.

How do we Further Improve Pain Management?
It appears that students entering medical school already have established viewpoints about pain and pain management.52 The current process of medical training may negatively reinforce attitudes toward pain management. With almost no education about pain in medical schools, attitudinal priorities are being established for diagnosis and treatment of disease, not symptom management. Medical schools expect to create physicians with a set body of knowledge but make little effort to account for pre-existing attitudes or cultural viewpoints. Individual psychological characteristics with respect to reluctance to prescribe opioids, and fears of patient addiction and drug regulatory agency sanctions are not addressed. If improvement of pain is to occur, then there needs to be modification of the teaching of pain and pain management at the medical school level, continuing on through residency training. JCAHO policy has ensured the documentation of pain, but it does not educate health care workers, correct attitudes or remove existing barriers.53 Despite linking such policy with accreditation, there is little evidence that such documentation has improved pain management and understanding in the ED. Skepticism of pain reporting will almost certainly nullify any value of documentation. It may be that more focused initiatives as seen with postoperative pain or back pain may

If improvement of pain is to occur, then there needs to be modification of the teaching of pain and pain management at the medical school level, continuing on through residency training.
be more successful than a single guideline that cannot be expected to be applicable due to the widespread origins and duration of pain seen in the ED.

Recognition of surgeon inability (or reluctance) to manage pain resulted in the establishment of acute pain services in almost all hospitals. Similar poor pain management for vaso-occlusive crises in EDs has resulted in sickle cell day hospitals. In that setting patients receive more rapid and more effective pain control than in the ED. They also are admitted less often and stay for shorter periods.54 Both solutions, while better for the patients, have not improved the pain management skills of those who were ineffectively managing patient pain. Removal of responsibility for pain management and failure to educate groups of physicians will probably result in other patients in their care continuing to receive suboptimal pain management. This has certainly been one of the reasons for the ongoing disagreement between surgeons and emergency physicians over pain control in patients presenting with abdominal pain.

Clinical practice guidelines appear able to provide at best a framework around which we can build individualized approaches to pain. While endpoints of pain control cannot be mandated, offer of pain control should be universal. Education about patient pain barriers will allow physicians to recognize and account for such barriers, ensuring the patient does not suffer needlessly. Furthermore physician attitudes toward self-reporting of pain, opiophobia and malingering need to be addressed. Only when these issues are addressed can steps to improve pain management be expected to gain widespread acceptance.

Further improvements will require policy that encourages research for areas lacking adequate evidence and more effective information dissemination strategies. Existing information could ensure better pain control in many aspects of emergency medicine if properly disseminated. This may require novel strategies given that current methods of disseminating information are for the most part unsuccessful. Until then, clinical guidelines will continue to be poorly used and minimally effective at modifying clinical practice. Improving pain management will thus be a slow process. It may well be that the greatest success will be achieved not from the top down, but from the bottom up: initiation of pain education early in medical school, with attention to student attitudes and biases. As these new physicians enter the clinical domain, they will pressure more senior physicians into modifying their attitudes and behavior, ultimately achieving optimal pain control for our patients.

Acknowledgement
The author gratefully acknowledges the support of the Mayday Fund in the development of this article.

References


29. See Jacobson, supra note 9, at 74H.

30. See Finder, supra note 27, at 69.


32. See Jacobson, supra note 9.

33. Ibid at 77H.


38. See Dobbins 2004, supra note 8.


44. See Jacobson, supra note 9.

45. See Gallagher, supra note 40.


49. See JCAHO, supra note 20.


51. See Wyatt, supra note 34.


53. See JCAHO, supra note 20.


55. See Wyatt, supra note 34.


57. See Finder, supra note 27.

O
ver the past thirty years, the doctrine of
informed consent has become a focal point in
discussions of medical ethics. The literature
of informed consent explores the evolution of the
principle of autonomy, purportedly emerging from
the mists of 19th Century medical practice, and find-
ing its earliest articulation in legal cases where
wronged citizens asserted their rights against medical
authority. A commonplace, if not obligatory, feature of
that literature is a reference to the case of Mary
Schloendorff and the opinion written by Judge
Benjamin Cardozo by which the case is remembered.
Commentators today applaud the prescience of
Cardozo for an early articulation of what eventually
would become bioethical orthodoxy concerning
informed consent and its place as a bulwark of patient
autonomy. They inevitably quote Cardozo's famous
statement, "Every human being of adult years and
sound mind has a right to determine what shall be
done with his own body."1 We should not make too
much of this sound-bite repetition of Cardozo's dic-
tum; it would be surprising to find a serious commen-
tator who used the Schloendorff opinion as the foun-
dation of an argument about the origins of informed
consent. Nevertheless, the quotation occurs often
enough in such arguments to make examining its
provenance worth the effort.

The Schloendorff case was brought to address a
claim of medical malpractice in which a surgeon was
accused of operating on an unwilling patient. We
know from the opinion that medical examination of
Mary Schloendorff revealed a tumor that her doctors
wished to examine more thoroughly “under ether.”
She consented to the ether, but claimed at trial that
she had withheld consent for surgery. By her account,
the doctors operated while she was unconscious and
despite her earlier objections. Schloendorff argued
later that she lost fingers to gangrene as the result of
the operation, and suffered injuries to her leg as well.

Consider how a leading text in biomedical ethics
uses Cardozo's opinion to trace the development of
autonomous choice through a chain of legal prece-
dents:

The best known and ultimately most influential of
these cases is Schloendorff v. New York Hospital
(1914). Schloendorff used rights of “self determina-
tion” to justify imposing an obligation to obtain a
patient's consent. Subsequent cases that followed
and relied upon Schloendorff implicitly adopted its
justificatory rationale.2

Other texts focusing specifically on the origins of
informed consent pay similar homage to the Schloendorff case, designating it a touchstone for all
manner of rights, including the right to bodily integri-
ty, the “sanctity of the person,” the right to refuse
treatment, and medical self-determination. In gener-
al, it has become a starting point for most discussions
of informed consent.3

Yet despite the attention given to Cardozo's opinion,
little has been written about the Schloendorff litiga-
tion itself. An exploration of the records from the case
might lead us to ask whether celebration of
Schloendorff is warranted. When we encounter the
Schloendorff opinion, are we applauding the first
appearance of patient's rights or merely reveling in
Judge Cardozo's penchant for phrasemaking? Did this

---

Phantom Tumors and Hysterical
Women: Revising our View of the
Schloendorff Case

Paul A. Lombardo

Paul A. Lombardo, Ph.D., J.D., is Director of the Program in
Law and Medicine in the University of Virginia's Center for
Biomedical Ethics. His research interests include the history
of medical litigation.
case actually signal a sea change in the relationship between doctors and patients?

It is rarely clear in most discussions of the Cardozo opinion that Mary Schloendorff lost her case. That result is not only startling because of the way Cardozo ignored the absence of consent for dangerous and unwanted surgery, but also for its extraordinary deference to charitable immunity of hospitals, employing questionable arguments and contorted interpretations of the facts to reach a conclusion that would allow the case to be dismissed. The very Court that Cardozo sat on – New York’s Court of Appeals – criticized the reasoning on charitable immunity in the Schloendorff case as “logically weak” only ten years after it was decided, and it was completely overruled in 1957 when the shield of non-profit status was discarded in New York as “out of tune with life about us.” Yet we still celebrate the case as a salute to patient autonomy.

It is instructive to view the narrative contained in the Cardozo opinion as a counterpoint to a contemporary news account as well as the story told by Mary Schloendorff during her testimony at trial. These contrasting treatments of what happened at the New York Hospital yield surprising insights about the practice of medicine, particularly surgery, the deference paid to the charitable work of hospitals, and the fate of patients – particularly women – who were the victims of medical malpractice in the first decade of the 20th Century. They also suggest that Cardozo was far from the godfather of informed consent some commentary would have us believe.

The Cardozo Opinion

The starting point in the Schloendorff story is the opinion of Benjamin Cardozo. His short legal essay is the tip of the historical iceberg – but the only truly visible point of reference for over ninety years since the case was decided. Cardozo ran successfully for a seat on the Supreme Court, a trial court in the New York State system, and took his place as a novice judge on January 5, 1914. Presiding over only six cases, he had barely tasted the life of a trial judge when New York Governor Martin Glynn appointed him to fill a vacancy on New York’s highest appellate tribunal, the Court of Appeals. Cardozo was on the bench less than six weeks when he heard oral arguments in the Schloendorff case, and he rendered the now famous opinion just over two months later, on April 14, 1914.

By then the case had been in the court system six years. Cardozo was forty-four years old, and his inexperience and status as a temporary judge (he was later elected for a full term on the Court in 1917) belies the notion that the Schloendorff opinion was noteworthy because of its author’s status. Cardozo’s reputation was made in the twenty-four years after Schloendorff, when he wrote dozens of memorable opinions, eventually becoming the Chief Judge of the New York court. It was this reputation that would eventually bring him to the U.S. Supreme Court in 1932, where he remained until his death in 1938.

Cardozo’s opinion begins with a focus on the single fact that will determine the case’s outcome: New York Hospital was a charitable institution and thus immune under state law from Schloendorff’s claim. From its opening lines, the opinion fairly reeks with deference for the role of medical philanthropy. Cardozo tells us that the hospital has a history stretching back to 1771, and a pedigree that includes a “royal charter” from none other than George III. Its raison detre is “the care and healing of the sick.” It is not a business enterprise, and has neither stock nor profits. Care and boarding are provided for free to those who need them; even paying patients are charged only seven dollars a week for lodging. It is, most certainly, a charitable institution.

Cardozo then turned to a terse description of the plaintiff, Mary Gamble Schloendorff, and her medical concerns. She arrived at the hospital in January, 1908, her complaint being “some disorder of the stomach.” She became a patient at the hospital, paying the charge for boarding of $7 a week. After several weeks a doctor there discovered a lump, which proved to be a fibroid tumor. Another physician advised an operation.

Cardozo reluctantly repeated what was apparently the most troubling result of the surgery – that “gangrene developed in her left arm, some of her fingers had to be amputated, and her sufferings were intense.” For purposes of discussing the legal claims that were made in the case, he allowed that this account must be tentatively accepted “even if improbable.”

Cardozo continued with a recitation of the legal posture of Schloendorff’s claim, in which “the wrong complained of is not merely negligence. It is trespass.” This was a critical point. Schloendorff was not claiming that she had suffered from the carelessness of the hospital and its surgeons. She charged that an operation on her was undertaken against her will and after she specifically stated her wishes to the contrary. Her
claim was couched in the language of battery, an intentional wrongdoing, recognized by the common law as offensive contact that yielded damage. Cardozo took the occasion of addressing the alleged battery to issue this ringing pronouncement on the right to self-determination in the medical realm.

Every human being of adult years and sound mind has a right to determine what shall be done with his own body; and a surgeon who performs an operation without his patient’s consent commits an assault, for which he is liable in damages.10

Cardozo left some room for exceptions to this sweeping rule, noting that “cases of emergency where the patient is unconscious” might justify operations without consent. But he emphasized the novelty of Schloendorff’s claim as separating it from the usual malpractice case, and he underscored the potential liability of the perpetrators of surgery without consent. He then established the legal distance between the hospital and its employees, asserting that the hospital, in making its facilities available for surgery, was not responsible for the potential misdeeds of its staff.

Again undercutting Mary Schloendorff’s account, he conceded that “if we are to credit the plaintiff’s narrative” it was the physicians who “ordered that an operation be performed on her in disregard of her instructions.” He separated the physicians from the hospital, as men serving their own profession, an “independent calling…sanctioned by a solemn oath and safeguarded by stringent penalties.” Whatever wrong may have happened, it may have been the doctor’s fault, but it was not the responsibility of the hospital.11

The opinion ended with the same bow to the role of medical philanthropy with which it began, endorsing the conclusion of the trial judge who directed the jury to reject the Schloendorff claim. It would be a shame, Cardozo declared, to rule against the hospital and thus “constrain charitable institutions” to limit their beneficence out of fear for liability.

The first assertion was based on “Good Samaritan” logic: if a charity patient who stood to benefit from the benevolent ministrations of a free hospital was negligently harmed in the midst of receiving care, he should not be allowed to sue the hospital. That line of argument became known as the “implied waiver” theory: by making yourself available as a beneficiary of charity, you waive your right to later turn on your benefactor.

The second basis for denying liability was the doctrine known as respondeat superior. That Latin term (“let the master answer”) refers to the legal responsibility assumed by employers when their employees (“servants” in the common law parlance) harmed someone else by failing to exercise adequate care while doing their jobs. The harmful acts of the employee were imputed to the employer (the “master”).

There were problems with each of these legal theories. Many people arriving at hospitals were unconscious, others incapable of understanding the legal conditions of their admission. The idea that they had implicitly “waived” their right to bring suit was too much a fiction for some courts to accept.16 As for respondeat superior, it was not always easy to determine whether harms were caused by independent non-employee doctors, or by unpaid interns, students, nurses, orderlies and the like. Some were clearly under the control of the hospitals, others less so.

Charitable Immunity

The legal doctrine that insulated charitable organizations from legal liability for torts has a spotty American pedigree. It first entered the U.S. through a somewhat questionable back door when the Supreme Court of Massachusetts embraced the concept in an 1876 case involving claims of a charity patient who suffered damages following surgery by a student-doctor.13 That court characterized the donations which supported the charity hospital as a trust that should be immune from certain damage claims. Authority for that proposition was not derived from American case law, but rather an English decision, Holliday v. St. Leonard’s,14 which declared charities should not be responsible for the negligence of their agents. Unfortunately, the Massachusetts court neglected to note that the Holliday precedent had been obliterated by several cases, at least one of which was decided a full ten years earlier, resoundingly overruling Holliday and rejecting its logic.15 Nevertheless, U.S. courts repeated the argument of the Massachusetts decision, which soon morphed into two distinct assertions about the tort liability of hospitals.

It would be a shame, Cardozo declared, to rule against the hospital and thus “constrain charitable institutions” to limit their beneficence out of fear for liability.
To complicate the picture further, almost as soon as charitable immunity rules began to be adopted in the U.S., the very notion of a pure “charity hospital” came under scrutiny. By the 1890’s, ostensibly “charity” hospitals had provisions for charging room and board to patients who could pay. Some hospitals also levied specific charges for professional services provided by doctors. These payments from financially capable patients were set aside to subsidize care for the indigent. The result was an awkward disconnect between the supposedly clear policy objectives of immunity as recited in state case law concerning hospitals, and the significantly more varied financial and administrative circumstances of hospitals and doctors.

The Schloendorff case occurred at a time when New York State courts had relied on both the “implied waiver” and respondeat superior theories to protect charitable hospitals. But the facts of the case made the use of either open to question. It was clear that at least as to room and board, Mrs. Schloendorff was a paying patient. It was also established that fees were paid to some physicians and surgeons who practiced at the hospital, and that those fees were sometimes collected by the hospital. More importantly, the Schloendorff case did not involve a charge of negligence, but trespass – and intentional tort. The New York cases on “implied waiver” covered negligence, not trespass.

Fitting the roles of doctors and nurses into the theory of respondeat superior was even more problematic. While the visiting surgeon who performed the operation on Mrs. Schloendorff seems to have been a truly independent contractor for whose behavior the hospital was not responsible, the nurses filled a very different position. They were, in fact, employees of the hospital and under its control. But Judge Cardozo’s opinion strained not only to make them answer only to the doctor, but to completely divorce their responsibility from that of the hospital.

Mrs. Schloendorff testified that she asked the nurses repeatedly about the “ether examination,” stressing that she had not consented to surgery. The implication of these comments was that the hospital – via its employees the nurses – was on notice of the pending operation was improper. The nurse is considered an automaton under the complete thrall of the physician, unable to discern when things might be amiss. Even though the testimony indicated that nurses had heard Schloendorff’s concerns about surgery, Cardozo imputes to them a “see no evil” simple-mindedness, which he attributes to both their role and their training.

Turning the role of the nurse as hospital employee on its head, Cardozo suggested that no one assisting in surgery will generate liability for the hospital despite the fact that a nurse may well have known the operation was improper. The nurse is considered an independent contractor under the complete control of the physician.

An ether examination was intended, and how soon the operation was to follow, if at all, the nurse had no means of knowing. Still less had she reason to suspect that it would follow against the plaintiff’s orders. About such matters a nurse is not qualified to judge. She is drilled to habits of strict obedience. She is accustomed to rely unquestioningly upon the judgment of her superiors. No woman occupying such a position would reasonably infer from the plaintiff’s words that it was the purpose of the surgeons to operate whether the plaintiff forbade it or not.

Cardozo admitted another potential exception to the immunity rule: “I can conceive” he suggested, “of cases where a patient’s struggles or outcries in the effort to avoid an operation might be such as to give notice” to hospital officials that some wrongdoing was afoot. But despite what Cardozo had learned from the trial record about Schloendorff’s cries of distress, this was no such case. The trial judge’s action – deciding the outcome without even submitting the case to a jury for a weighing of the facts – was allowed to stand. Cardozo’s biographer noted how restrictive Cardozo’s view of medical liability was in Schloendorff, which represented a “refusal to alter rules that exempted employers from liability.” “Not only did he accept and extend the immunity of a charitable hospital from liability for the negligence of physicians and nurses using its facilities, but he also extended the immunity to cover negligence of the hospital’s teaching staff.”

The opinion, which sounds in part so sympathetic to the plight of victims who face medical abuse, is in its result extremely protective of medical prerogatives. The loss of a few fingers, however tragic, is not enough to justify changing the rule of charitable immunity; the critique of hypothetical “ministers of healing… [who] proved unfaithful to their trust” will not apply to the physicians at the New York Hospital, whose “solemn oath” has apparently not been breached.

This was, in fact, a very conservative opinion, yet one whose most memorable lines sound, in hindsight,
expansively sympathetic to patients suing doctors. Its description of Mary Schloendorff’s surgery and its aftermath provides a contrast to other accounts, including one that was supplied by New York’s press.

**Schloendorff in the News**

The contemporary public report of the *Schloendorff* controversy could not have been less sympathetic to the plight of beleaguered hospitals, nor more attentive to the alleged harms of Mary Schloendorff. *The New York World* carried this headline, announcing the beginning of the trial against the hospital where Mary Schloendorff endured surprise surgery.

**HER PHANTOM TUMOR LEFT HER ARM USELESS**

**MRS. SCHLOENDORFF ASKS $50,000 DAMAGES FOR OPERATION**

The story that followed introduced readers to the newlywed Mary Schloendorff. She was, said the paper, formerly an elocution teacher in San Francisco known as Mary Gamble. She claimed that surgery was performed on her without her consent and, as a result, she was “maimed for life.” At trial in a New York courtroom, she told her story. The artist’s courtroom sketch of Mrs. Schloendorff showed an elegant and dignified lady on the witness stand. Having survived the San Francisco earthquake in 1906, she suffered from “nervous shock.” After entering the hospital in January 1907 she was prepared to leave when “some of the medical staff suggested that she undergo an examination” to determine her “exact physical condition.” Mrs. Schloendorff “had a lump in her side” for five years and “of which her nurses knew.” A doctor suggested an “ether examination” saying that she was too nervous to examine her otherwise.

To this she also consented, but stipulated that there be no operation at that time. The surgeons called the lump a “phantom tumor” and gave her the impression that it might be due to nervousness. On recovery from the effects of the ether, Mrs. Schloendorff found that the surgeons had made incisions in her back and abdomen. In consequence, she testified, her fingers developed gangrene and she lost the use of her right arm.

The paper reported that the hospital denied all responsibility, claiming that “the operation was properly and skillfully performed, and that afterward the patient received proper care, attention and treatment.” It did not report the result of the trial, cut short after four days when the judge directed the jury to render a verdict absolving the hospital of liability. Court costs of $292.62 were charged to Mrs. Schloendorff. The case was appealed, but the first court to review the trial result upheld the decision against Schloendorff.

**Mary Schloendorff’s Account**

The transcript of the *Schloendorff* trial provides us with a dramatic contrast to both the severely truncated, and at times hostile, narrative contained in the Cardozo opinion, and also the newspaper story that summarized the trial. Those other accounts masked what the records of the case reveal, unconstrained by the delicate conventions of the early 1900s. Mary Schloendorff had no ordinary tumor, nor was it on her arm. It was a fibroid mass in her uterus and the operation done against her will was a hysterectomy.

Mary Schloendorff introduced herself as the first witness at trial with the following statement.

I lived prior to November, 1906 in the City of San Francisco, California. I lived there nearly all my life. I was a teacher of physical training, voice, and culture, of reduction and development. My physical condition in the fall of 1906 was, I might say, perfect. The earthquake of San Francisco occurred on April 18, 1906. Well, I was greatly frightened and nervous, of course.

This admission of “nervousness” following the great earthquake was reiterated by several later witnesses, not as the explanation for Schloendorff’s stomach problems, but as the basis for arguing that she was noncompliant as a patient and unreliable as a witness.

Fleeing her memories of the earthquake, she arrived in New York City in September of 1906, and lived with her adult son for a month before taking up residence in a rooming house. She sought medical advice from a doctor for what she characterized as “dyspepsia or indigestion.” She took “Stewart’s dyspeptic tablets and Bromo Seltzer” to alleviate her pain. In early December of 1907 she went to New York Hospital, which her physician had recommended. There she was treated by a Dr. Frederick H. Bartlett, who after approximately a month, told her she was “cured.” “He told me I could go home at the end of the second week,” she said, “but if I preferred to stay a week longer and gain strength, I could, as I was very much reduced in flesh, and I concluded to stay the week out.”

But before she was discharged, Dr. Bartlett performed a physical examination that revealed a lump. This was no news to Mrs. Schloendorff, who stated that the lump in her side had been evident for about
five years. Bartlett asked to bring in his colleague Dr. Lewis A. Stimson for a consult, but upon a second examination the lump had disappeared. Bartlett said it might be a floating kidney or “phantom tumor.” Stimson (the surgeon) attributed his inability to locate the tumor to Mary Schloendorff’s demeanor. He said that she was “too nervous, too rigid or too tense.” He recommended an “ether examination” so that the doctors could locate the mass. Schloendorff was cautious about the nature of this examination.

But I asked Dr. Bartlett the next time that I saw him what was meant by an ether examination, and he said it meant to give the patient a little ether to quiet the nerves and relax the body. That it didn’t amount to anything; and I told him “I don’t want any operation Doctor.”

Dr. Bartlett advised her to see the surgeon while she was in the hospital to determine what the lump was. If it required surgery, she could return in the future. Schloendorff testified that her mind was made up to go home. She packed her suitcase and wrote a letter to her landlady saying that she was returning. The landlady would later confirm this sequence of events as part of her own testimony, and would introduce the letter into evidence.

But the night before Mary Schloendorff was to leave she was awakened by the night nurse who moved her to another ward. Again, toward midnight yet another nurse awakened her to prepare for the “ether examination.” Mrs. Schloendorff stated several times that she wanted no operation, but was reassured that only the “ether examination” would take place.

She was “swathed in antiseptic cloths...tied up like a mummy and placed back on the bed” to sleep. Upon rising the next morning, she was wheeled to another room where a man prepared to give her “gas” as a preparation for delivering the ether. She struggled, and tried to leave, but was restrained.

He had some apparatus there with a rubber tube and mouthpiece, and he took his hand and pushed against my forehead and pushed me back and put the mouthpiece to my mouth and said “Take a deep breath.” I was frightened at the gas and tried to get up, took a deep breath, I guess, and did not know any more.

When she regained consciousness, there was a large scar and pain. There was no contact with the surgeon for two weeks, nor was there any explanation of the nature of the surgery. During the lawsuit the surgeon eventually testified how in order to remove the tumor, he had to cut into the abdomen and tie off four arteries. All this was in preparation for the most critical part of the operation:

I...cut off the uterus at the junction of the neck and the body and took out that part; took out the upper portion. I took out all except the neck of the womb. I did take out the ovaries. I did not take out anything else.

He described the procedure in medical terms as “removal of the uterus, which was the seat of a fibromyoma, technical name of the operation being supravaginal hysterectomy.”

Mrs. Schloendorff reported strange occurrences following surgery. She had been moved to the basement to hide her cries of pain. From the colloquy between her lawyer and the nurses on duty, removal to this setting may have led to earlier legal actions by other patients who felt abused. Schloendorff testified that the night nurse ordered orderlies to carry her to the basement. On her third night there, a woman arrived from an ambulance in the street. She was “ragged and covered with blood.” Schloendorff observed this woman from her perspective “lashed to the bed.”

Schloendorff’s lawyers tried unsuccessfully to introduce evidence that orderlies had threatened “to break her neck” if she didn’t muffle her cries. A nurse noted that Schloendorff “was keeping the other patients awake. She was in great pain.” The lawyers tried to suggest that the nurse had been arrested for assaulting another patient, but the questioning was not allowed. Schloendorff’s testimony concerning her pain was particularly graphic.

I was cut across the stomach from hip to hip. I suffered a great deal, more than tongue can tell...My mouth was torn to pieces inside...I suffered with pains in my arm and hand, coldness and numbness in my left hand.

But the responses of her doctors generally minimized her suffering, treating it as unlikely or imagined. Her surgeon believed that “it did not amount to anything, it would pass away.” Though Schloendorff recalled that her “hand was cold and the nails were blue;” the doctor “laughed and said that I was very imaginative.”

Upon making his rounds Dr. Stimson (the surgeon) “playfully punched me with his fist in the abdomen; in a playful way he said, ‘How are you, how do you feel old girl?’ I screamed in agony at the pain he gave me.” The same explanations for her “imagined” pain were given to her son Evan Gamble. “[S]he was suf-
ferring no pain” the doctor said, “she only imagined it.” But the problem with her hand continued. She said her “left hand is useless. Fingers turned back and the cords – permanent adhesion, permanently turned back….One of my legs is affected almost as much as my hand.”

What possible association could there be between a hysterectomy, consent or not, and the withered hand that eventually brought Mary Schloendorff to court? A physician retained as an expert witness in the trial connected the two events. Dr. George Schoeps treated Mrs. Schloendorff from 1909 to 1911, eventually billing her for over $100. He testified that an embolism formed as a result of the surgery caused the condition of her hand. He told the court:

An embolus is a foreign body in the circulation...A piece of fat, of coagulated blood or a piece of tissue. ...Now it is possible from this, and very likely from the wound surface, that some blood or foreign material went into the circulation and in the course of the blood circulated through the arteries and the heart...causing endocarditis.

Schoeps declared that surgery was not always necessary, and that medical treatment was available for fibroid tumors. He also noted that an examination of such tumors could be done without using ether. The treating doctors responded that the problem with Mrs. Schloendorff’s hand was her own fault. “She did not keep her hand elevated. The direction was not followed.” Another doctor said that “…this arm was kept as quiet as possible, for a long period, by means of bandages, dressings, splints, and pillows, no one of which means could be used continually because of the patient’s extreme nervousness and impatience.”

At the time of Mrs. Schloendorff’s operation in 1907, surgery for fibroid tumors in women was a somewhat recent development. The first successful total hysterectomy for fibroid tumors was reported in New York in 1888. Other available treatments included various medications, and the process of passing an electrical current through the uterus. The latter was a controversial but nevertheless common treatment for a while, purportedly used to shrink the tumor. There was also always the choice not to treat the tumor at all, and some practitioners suggested that women with no other symptoms than the mass itself “had better refuse to be operated on.”

Regardless of the efficacy of therapeutic alternatives, the danger of surgery motivated many surgeons to urge caution in elective cases. While some came down soundly in favor of operating, others made it clear that it was the patient’s prerogative to decline the operation. Advising surgery was a “serious matter” when the mortality rate from operating ran as high as eight percent, and though the rate had dropped measurably by the time Mrs. Schloendorff’s case was decided, medical debate over the proper indications for surgery provided “a perennial source of strife” among surgeons.

Approximately one month before she finally left New York Hospital, Mary Schloendorff was prepared for discharge. “I was told that I was cured” she said. “And they dragged me down, went through the formalities of a discharge from the hospital.” Upon leaving the building, she collapsed in the street. Confronted with her condition, one of the doctors reportedly said “that was all a mistake; we did not intend for you to go.” She remembered subsequently staying between five and six weeks more.

Under orders of the hospital management, Schloendorff was discharged in April of 1907 to St. Andrews convalescent home, where she stayed only briefly. She then moved to New York Graduate Hospital (two weeks) then to French Hospital (approximately three months). She was admitted to Bellevue Hospital for an operation on her hand, (two-three months) and then to Lebanon Hospital (five weeks) for convalescence the next summer. She was also treated at Cornell Hospital. At Bellevue she used her maiden name, Berry, “because I wished to hide myself from the world where no one would find me.” This attempted anonymity was later used as evidence to impeach her credibility as a witness, and undermine the rest of her testimony as unreliable.

Others testifying on Schloendorff’s behalf supported her recollections, one witness charging that a young surgeon described the botched operation as “all a terrible mistake.”

Consent

Mary Schloendorff’s prime contention, that surgery had occurred without consent, was vigorously contested by the defense. Dr. Lewis Stimson, the operating surgeon, stated that he kept no records of opera-
tions.\textsuperscript{60} But he recalled that his encounter with Mrs. Schloendorff consisted of a manual examination of the tumor and these comments: “I told her I would remove it if she wanted it removed. She did not say she was opposed to an operation.”\textsuperscript{61} Upon cross-examination he reiterated: “I told her if she wanted to have the tumor removed she could come over on the surgical side and I would take it out.”\textsuperscript{62}

There were no other records to review on the question of consent because, as Dr. Bartlett testified, “I do not believe it is a custom...to give consent in writing to an operation.”\textsuperscript{63} Schloendorff’s son Evan Gamble testified that he had been assured that no operation would be performed without notification; that was the custom in the hospital. Gamble insisted that he “did not wish her to be operated on even with her consent, that mine had to be obtained.”\textsuperscript{64} Dr. Stimson asserted that he would “never operate on a person without their consent.” But neither would he consult with the relative who brought an adult patient to the hospital.\textsuperscript{65} Like Dr. Bartlett, he declared that it was not the custom to consent in writing at the New York Hospital.\textsuperscript{66}

What was the usual custom concerning consent in 1914? As scholars have repeatedly pointed out, while traditional medical codes said little about consent, by the early Twentieth Century the expectation that the consent of patients was required before treatment was well settled. Particularly in cases involving surgery, patient wishes were usually followed.\textsuperscript{67} By the time of Cardozo’s Schloendorff opinion, the law was also clear in many jurisdictions, as Cardozo well knew, since he cited two such cases in the Schloendorff opinion.

One was the 1905 Minnesota Supreme Court decision in \textit{Mohr v. Williams}. A trial jury awarded over $14,000 in damages to Anna Mohr for the loss of hearing in her left ear following an operation by Dr. Cornelius Williams. She contended that she had only hearing in her left ear following an operation by Dr. Cornelius Williams. She contended that she had only

Echoing this perspective, the court thought the requirement of consent to be such a settled matter that “[i]t cannot be doubted that ordinarily the patient must be consulted, and his consent given, before a physician may operate on him.”\textsuperscript{69} The choice of proper treatments, and the methods of delivery, were matters for a doctor’s judgment, but the court could find no legal principle that would give surgeons “free license” to operate.\textsuperscript{70} Moreover, a patient, like any other person, has a right to expect “complete immunity of his person from physical interference of others.” The court characterized an unconsented surgery as a “violent assault.”\textsuperscript{71}

The second case cited by Cardozo was the 1906 Illinois case of \textit{Pratt v. Davis} involving a woman who, like Mary Schloendorff, endured a hysterectomy. Parmelia J. Davis suffered from epilepsy and went to a sanitarium for treatment twice in 1896. Dr. Edwin H. Pratt provided “minor surgical treatment” during her first visit, but the second time, without any disclosure of his intentions, had her anesthetized with chloroform and surgically removed her uterus. Pratt testified that he intentionally deceived Mrs. Davis so that she would comply with the operation. The doctor excused his actions with the assertion that she was insane (she was, at the time of trial, committed to an asylum) and the surgery was a treatment for her condition. Additionally, he said he had a general consent from her husband to do whatever was in her best interest, and “implicit” consent from her since she was a voluntary patient at his facility.

Mr. Davis responded with a lawsuit for malpractice, using the theory that surgery without consent, regardless of therapeutic motive, created liability for “trespass” and justified damages. The suit resulted in a judgment against the doctor for $3000 which doctor Pratt challenged on appeal to the Illinois Supreme Court. Ruling against Pratt, that court said that “...where the patient is in full possession of all his mental faculties...it is manifest that his consent should be a prerequisite to a surgical operation.”\textsuperscript{72} Consent from someone with legal authority – spouse, family or others – was necessary for surgery even for mentally incompetent patients, except in cases of “great emergency.”

Similar reasoning could be found in other cases, such as the 1913 Texas decision of \textit{Rishworth v. Moss}, which not only required consent for surgery, but was very specific about who could provide it validly. Eleven year old Imogen Rishworth was taken to see Dr. Robert E. Moss by her sister, an adult. The sister gave consent for surgery to remove her tonsils, but the child died after an administration of chloroform for anesthetic purposes. The lawsuit, subsequently brought by Imogen’s parents, was dismissed following an instruction by the trial judge who told the jury that consent of the adult daughter was equivalent to con-
sent of the parents. But no evidence was presented suggesting that the daughter had the legal right to consent, even for her own sister.

On appeal, the Texas court declared that "...it seems to be reasonably established that a physician is liable for operating upon a patient unless he obtains the consent of the patient, if competent, and if not, of some one, who, under the circumstances, would be legally authorized to give the requisite consent." An instruction that took away the legal authority of the parents to give or withhold consent was "absolutely and fundamentally wrong." The 1913 Oklahoma surgical case of Rolater v. Strain was decided similarly. Mattie Strain stepped on a nail, and went to Dr. J. B. Rolater to have the subsequent inflammation drained. During this process the doctor found a bone he judged out of place, and removed it. Strain sued, claiming she had specifically asked that no bones be removed. Rolater said that removal was necessary as an emergency matter but the court disagreed, finding no emergency that would justify surgery directly contrary to the patient's stated wishes. It concluded that removing a bone without consent of the patient, "was therefore unlawful and wrongful, and constituted a trespass upon her person." Rolater, like the cases cited by Cardozo, points to the same conclusion: At the time of Schloendorff, the rules were reasonably clear that Doctors were expected to get consent before surgery both as a matter of medical custom and the law.

Conclusion

What then do we learn from Mary Schloendorff's account at trial in contrast to the Cardozo opinion? Despite the protests of hospital officials and their lawyers, and based on the literal testimony of Dr. Stimson the surgeon, it would appear that Mary Schloendorff gave no explicit permission for an operation. However else we credit her testimony, the surgeon's declaration reveals that the only defense he could muster was that she had not directly refused his services. According to him she never said yes, she just didn't say no. So while this might be a case about the need for simple consent, it is hardly a case about informed consent. At best Cardozo's opinion merely restates a maxim that could even at the time of trial be traced back through two hundred and fifty years in American law: surgery without consent is actionable; in some instances, it could even be considered a crime.

Consent was undoubtedly required of surgeons in the Schloendorff era, and the physicians who testified in Mary Schloendorff's trial said nothing to contradict that expectation. But the theory of informed consent, requiring an explanation of risks, benefits and alternatives to aid patient understanding and honor patient autonomy, took form only haltingly after World War II, and was nowhere to be seen in 1914 medical jurisprudence. The Schloendorff case is more properly read as a restatement of a none-too sturdy doctrine of charitable immunity.

Schloendorff is a reminder of the power of a judge's rhetoric when taken out of context. Cardozo's opinion is certainly not the paean to personal freedom it is cited so often to represent. Even on its own terms it dismisses the very person whose case cries out for a remedy at law. Cardozo gives one brief shining epigram that catches our imagination: "Every human being of adult years and sound mind has a right to determine what shall be done with his own body." But to Mary Schloendorff this was less than an empty promise. Applied to her the opinion might more properly read: "Abandon all hope of justice, Ye who enter here." As this analysis shows, Cardozo's reputation as the godfather of informed consent rests on a very slender reed.

Finally, what does the case say about the role of Schloendorff in medico-legal history? The case should be famous not for Cardozo's opinion, but instead, for the fact that he never mentioned the major harm done to Mary Schloendorff. Her complaint was not merely that she lost fingers and the use of a hand; she endured a hysterectomy she did not want and probably did not need. Her pains were called "imaginary" and her troubles described as the product of "nervousness." The court's decision in favor of the New York Hospital on grounds of charitable immunity cut against what by 1914 was an already well-established sensitivity toward the right to medical choice. Schloendorff is a monument to the power of a judge's
rhetoric; notwithstanding its place in the bioethics literature, it does not merit its current reputation as a progressive salute to autonomy and a milestone on the road toward informed consent. Despite Cardozo's articulation of the right of medical self-determination that made her name famous, Mary Schloendorff's claim was rejected and the full story of the wrongs she may actually have suffered was lost along the way.

References
8. Schloendorff v. New York Hospital, supra note 1, at 127-128.
9. Id.
10. Schloendorff v. New York Hospital, supra note 1, at 129-130.
12. Schloendorff v. New York Hospital, supra note 1, at 135.
16. See Bobbe, supra note 13, for cases rejecting the “implied waiver” rule.
17. M. J. Vogel, The Invention of the Modern Hospital: Boston, 1870-1930 (Chicago and London: University of Chicago Press, 1960): 105-107. One commentator noted that long before Schloendorff states such as Rhode Island had concluded that even charitable hospitals could be held liable for the misadventures of surgeons, even though they received no pay from the hospital. E. B. Kinkead, Commentaries on the Law of Torts, vol 1 (San Francisco: Bancroft-Whitney, 1903): at 209-212. The Rhode Island case (Glavin v. Rhode Island Hospital, 12 R.I. 411 [1879]) was cited by Cardozo for a different proposition, to demonstrate that a master/servant relationship does not usually exist between outside doctors and hospitals.
18. Court of Appeals, State of New York, Briefs and Records, Mary E. Schloendorff against The Society of the New York Hospital, (1914) 1, Testimony of Henry Crane, at 176-177 (hereafter, Transcript).
20. Schloendorff v. New York Hospital, supra note 1, at 134.
23. Transcript, supra note 18, at 2.
24. Schloendorff v. Society of New York Hospital, 133 N.Y.S. 1143 (1912). Though the decision of the Appellate Division of the Supreme Court was rendered without an opinion, one judge noted his dissent.
26. Recurring references to Schloendorff’s “nervousness” suggest a barely submerged theory of hysteria linking parts of the medical testimony. It was not uncommon for doctors to claim that surgical interventions such as hysterectomy would cure a woman’s insanity. See, for example, “Insanity and Pelvic Lesions,” Medical News 76, February 3, 1900, at 176. For a summary of the literature on hysteria, see M. Micale, Approaching Hysteria: Disease and Its Interpretations (Princeton, NJ: Princeton University Press, 1995).
27. Transcript, supra note 18, at 17-19.
28. Ibid.
29. Transcript, supra note 18, at 20.
30. Ibid.
31. Transcript, supra note 18, at 21.
32. Transcript, supra note 18, at 22.
33. Transcript, supra note 18, at 192; Transcript, supra note 18, “Plaintiff’s Exhibit # 4,” January 30, 1907 (Letter to Emily Lux), at 205.
34. Transcript, supra note 18, at 24.
35. Transcript, supra note 18, at 35.
36. Transcript, supra note 18, at 124.
37. Transcript, supra note 18, at 54.
38. Transcript, supra note 18, at 169.
39. Transcript, supra note 18, at 163.
41. Transcript, supra note 18, at 196.
42. Transcript, supra note 18, at 51 (Testimony of Evan Gamble).
43. Transcript, supra note 18, at 31.
44. Transcript, supra note 18, at 56-61.
45. Transcript, supra note 18, at 62.
46. Transcript, supra note 18, at 67.
47. Transcript, supra note 18, at 149 (Testimony of Dr. Otto Goehle).
48. Transcript, supra note 18, at 189 (Testimony of Dr. George M. Cottle).
51. J. R. Goffe, “What Advice Should be Given to a Woman Suffering from Fibroid Tumor of the Uterus?” Medical News 82 (February 7, 1903): 247-249.
53. C. J. Webster, “A Consideration of Fibroid Tumors of the Uterus Based Upon a Series of Cases Treated Surgically,” Medical News 86 (April 22, 1905): 764-768.
55. Transcript, supra note 18, at 57.
56. Transcript, supra note 18, at 38.
57. Transcript, supra note 18, at 40-44.
58. Transcript, supra note 18, at 45 (Testimony of Lillias M. Reeve).
59. Transcript, supra note 18, at 191, 192 (Testimony of George M. Cottle and Testimony of Lillias M. Reeve, in rebuttal).
60. Transcript, supra note 18, at 125.
61. Transcript, supra note 18, at 110.
62. Transcript, supra note 18, at 117.
63. Transcript, supra note 18, at 108.
64. Transcript, supra note 18, at 49.
65. Transcript, supra note 18, at 129.
66. Let this issue seem too clear, yet another physician testified: “It is not the custom to operate without the consent of friends of the patient, such friends being matters of record in the office of the institution where they happen to be placed.” Transcript, supra note 18, at 167 (Testimony of Dr. William M. Polk).
69. Transcript, supra note 18, at 117.

69. Mohr v. Williams (95 Minn. 261, 268 [1905]).
70. Mohr v. Williams, supra note 69, at 269.
71. Mohr v. Williams, supra note 69, at 271.
72. Pratt v. Davis (224 Ill. 300, 305 [1906]).
73. Rishworth v. Moss (159 S. W. 122, 124 [Tex. 1913]).
74. Rishworth v. Moss, supra note 73, at 124.
75. Rolater v. Strain (137 Pac. 96, 98 [Okl. 1913]).

76. This is the conclusion of some of the most thorough commentary on the case as well; see J. Katz, The Silent World of Doctor and Patient (New York: Free Press, 1984): 51-52. Katz is the only commentator I have identified in the bioethics literature who realized that the operation on Mary Schloendorff was a hysterectomy. His attention to the gynecological focus of the Schloendorff surgery is echoed by Kathleen E. Powderly, who also challenges Katz’s view that consent was not commonly sought, while crediting the Schloendorff case with establishing the legal doctrine of informed consent. See K. E. Powderly, “Patient Consent and Negotiation in the Brooklyn Gynecological Practice of Alexander J. C. Skene: 1863-1900,” Journal of Medicine and Philosophy 25 (2000): 12-27. One recent Cardozo biographer also identifies the hysterectomy amid an extensive analysis of the case, describing it along with similar personal injury cases that generated Cardozo opinions as an example of judicial deference to the learned professions, see R. Polenberg, The World of Benjamin Cardozo (Cambridge, MA: Harvard University Press, 1997): 108-114.
Discourse on Embryo Science and Human Cloning in the United States and Great Britain: 1984-2002

Matthew Weed

There is a stark difference between American and British policy on embryo science and research cloning. The following survey of the discourse offered both in support of and in opposition to research cloning and embryo science in the United States and Great Britain will show that the same arguments were made in both countries. The fact that similar ethical argumentation occurred in environments where different policy was set is an indicator that current frames for ethical discourse on embryonic stem cell research and human cloning do not effectively capture the debate in the form that politicians and possible consumers of services to be derived from embryo science face.

The ethics surrounding embryo research and human cloning have been presented from virtually every possible viewpoint in all forms of medium. It is impossible to reprise every argument made on embryo science and research cloning; therefore, this survey will focus on some of the arguments made during the time leading up to the enactment of Great Britain's Human Fertilisation and Embryology Act of 1990 and the Human Fertilisation and Embryology regulations added to it in 2001. The cases for and against embryo research and nuclear transfer experiments in Great Britain and The United States will be described with particular attention paid to the debates over whether embryo research is acceptable, what benefits and dangers are attached to it, and at what point research should not be carried out because the embryo has developed sufficiently to require consideration as a human being.

This review will demonstrate the similarity of the arguments made in the United States and Great Britain on the use of embryos in science. Although the ethical literature makes up an important part of this paper, much of it will be dominated by reviews of political discourse on embryo research as represented by the Report of the British Committee of Inquiry into Human Fertilisation and Embryology in 1984; the debate in the House of Lords over the Human Fertilisation and Embryology bill in January of 2001 and the 2002 report of the President’s Council on Bioethics, Human Cloning and Human Dignity: An Ethical Inquiry.

Important Ethical Reviews, Policy Reports and Legislation

The report of the President’s Council on Bioethics (PCB), Human Cloning and Human Dignity: An Ethical Inquiry, is the current major United States document specifically focused on the ethics and policy of research cloning. It is important to note here that the President’s Council on Bioethics was asked to generate majority and minority reports on matters on which it could not reach consensus. In its July 2002 report, the ten-member majority of the PCB offered the view that a four-year moratorium should be placed on all human cloning research in the United States. The PCB also presented the views of a seven-member minority who believed that reproductive cloning should be banned but some form of regulated creation of embryonic stem cells via cell nuclear transfer (CNT) should be permitted.

Eighteen years before Human Cloning and Human Dignity: An Ethical Inquiry was published, the British Government received a report from Mary

Matthew Weed, Ph.D., is a visiting researcher in the Department of Genetics of Yale University, where he received his Ph.D. in 2004. He has also received a master’s degree in Public Affairs from Princeton University, which he received from its Woodrow Wilson School of Public and International Affairs in May of 1995. He has also received a Master of Arts Degree in Genetics from Harvard University.
Warnock, who was chair of the Committee of Inquiry into Human Fertilisation and Embryology. This committee, often called the “Warnock Committee,” set forth a variety of recommendations concerning the use of embryos in research. The Warnock Committee suggested how, when, and where experiments using human embryos might be carried out and proposed limits on clinical assisted reproduction in Great Britain. Five and a half years later, many of these recommendations became law when Parliament passed the Human Fertilisation and Embryology Act of 1990. The debate over whether embryo research should be permitted began years before the Warnock Committee report; likewise, societal discussions on the ethics of creating human embryos using CNT have not ended with publication of Human Cloning and Human Dignity: An Ethical Inquiry. However, these documents frame the time period in which British policy was established and grew under the auspices of the Human Fertilisation and Embryology Authority. They also frame much of the American political debate over whether embryo science and research cloning would be funded by the Federal government and, in fact, whether nuclear transfer experiments of any kind would be permitted in the United States.

Presenting the Cases in the United States and Great Britain

This article will use a topical approach with significant respect for the chronology of the debate in both countries to suggest that differences between American and British ethical discourse do not account for the difference in policy on embryo science between the United States and Great Britain. The order of arguments put forward in Human Cloning and Human Dignity will be a partial frame for the presentation of material here as it is the most recent government-based document in the United States and Great Britain that covers all of the ethical issues in notable depth. Their antecedents, whether rising in Britain or America, will be offered before each section is closed with the statements made in the report of the President’s Council on Bioethics.

The Case for Embryo Science and Research Cloning in the United States and Great Britain from 1984 to 2002

Those who support embryo research in both the United States and Great Britain alike believe that experiments using embryos, whether created using in vitro fertilization or CNT are valuable because they might be used to treat diseases. After deciding that embryo research is necessary, they have had to decide at what point in embryonic development it becomes ethically unacceptable to carry out experiments and debate the method of regulation (if any) under which this research should be carried out.

One of the major reasons to do embryo research of any kind is the hope that it will result in treatments for a wide variety of disease and infirmity. Chapter eleven of the Warnock Committee’s report made clear the belief of many British ethicists, scientists and others that assisted reproduction and research that might improve it were needed in order to treat human infertility.

...we are bound to take account of the fact that the advances in the treatment of infertility, which we have discussed in the earlier part of this report, could not have taken place without such research; and that continued research is essential, if advances in treatment and medical knowledge are to continue.

In 1999, President Clinton’s National Bioethics Advisory Commission (NBAC) discussed federal support for embryo research in the United States. Although it believed that the science should be regulated, the NBAC was clear that some form of regulated embryo research is ethically acceptable and scientifically necessary:

A principal ethical justification for public sponsorship of research with human ES [embryonic stem] or EG [embryonic germ] cells is that this research has the potential to produce health benefits for those who are suffering from serious and often fatal diseases.

Another part of the debate over the possible benefits of embryo research deals with the scientific knowledge that could be drawn from doing it. Lanza, et al., published an article in the Journal of the American Medical Association that expressed the hope that not only could embryo research cure disease directly, but that it might also lead to knowledge that would obviate the need for it altogether.

In the longer run, CNRT may prove to be only a transitional technology that is replaced by adult stem cell transdifferentiation. But far from devaluing cell activation research, this transitional role renders it even more important in the immediate future.

According to Lord Patel, a physician and chair of the Genetics Advisory Committee to Britain’s Medical Research Council, research to be allowed under the
proposed Human Fertilisation and Embryology regulations of 2001 might also facilitate this dream.12

Understanding how cell nuclear replacement into an unfertilized egg can reprogramme an adult cell nucleus and develop pluripotential stem cells...is necessary if scientists are to learn how to deregulate adult cells – any adult cell – to become pluripotential stem cells. When they do, there will no longer be a need to use human embryonic stem cells or the cells obtained following nuclear transfer.13

Writing in July of 2002, seven of the seventeen members of the President’s Council on Bioethics stated their belief that the creation of embryos using nuclear transfer was ethically acceptable and claimed that there were a variety of reasons why it should be done. As was the case with the Warnock Committee, National Bioethics Advisory Commission and in Parliament, these reasons focused on developing an understanding of the mechanisms of human development and the hope for cures for diseases that affect millions world wide.

The moral case for cloning-for-biomedical-research can be stated in the following straightforward way: American society and human communities in general have an obligation to try to heal the sick and relieve their suffering. This obligation, deeply rooted in the moral teaching of “love of neighbor,” lies heaviest on physicians and healthcare professionals who attend to individual patients. But it guides also the activities of biomedical scientists and biotechnologists whose pioneering research and discoveries provide new and better means of healing and relieving those who suffer. Research on cloned human embryos is one more path to discovering such means.14

One of the hopes that was expressed on both sides of the Atlantic ocean was that “cloning to produce immune-compatible tissues for transplantation” could become a reality. Speaking in debate on January 22, 2001, Lord Patel argued that this research would support:

...understanding how cell nuclear replacement into an unfertilized egg can reprogramme an adult cell nucleus and develop pluripotential stem cells....programmed to develop into different cell types and...used for transplant without the problems of host rejection....initial work will be done in mouse models, but it will need to be confirmed in humans.15

It was this hope that supported the argument of Lanza, et al., as quoted above. In its most generalized form, the belief that embryo research would lead to the curing of disease using whatever ethically acceptable means could be created was the most powerful reason that supporters of embryonic stem cell research had for their work.

However, they had to deal with the ethically thorny problem of when it was acceptable to use human embryos and when it was not. A key part of this question focused on the point in development beyond which doing research on an embryo became morally unacceptable, because it had developed sufficiently to be considered a human being and therefore not usable in research that would destroy it.

The Moral Status of the Embryo and Research Cloning

The 1984 report of the Warnock Committee in Great Britain recognized that there were many factors of which any policy on embryo research must take account. The view of the majority of the Warnock Committee that embryo science must be possible in order to develop new treatments for infertility is quoted above, and is the philosophical basis for the research provisions of the Human Fertilization and Embryology Act of 1990.

On January 22, 2001, with the Human Fertilisation and Embryology Authority already in place, the Bishop of Oxford stated:

Although the early embryo does not have an absolute status...it has nevertheless a special status. Research on it to produce stem cells could be justified morally only if this was the only way of obtaining the desired knowledge and if the knowledge was of such benefit that it outweighed that special status.16

Writing in July of 2002, seven members of the President’s Council on Bioethics expressed their belief that research using nuclear transfer derived embryos was ethically acceptable by arguing that:

...the fundamental moral judgment about whether to proceed with cloning-for-biomedical-research must be grounded in our judgment about the moral status of the embryos themselves, not the purpose of their creation. If an embryo or a cloned embryo had no moral standing, then creation for research and eventual destruction would present no moral problem. If the embryo or cloned embryo were morally the equivalent of a child, then regardless of how or why it was produced, experiments...
In its most generalized form, the belief that embryo research would lead to the curing of disease using whatever ethically acceptable means could be created was the most powerful reason that supporters of embryonic stem cell research had for their work.

upon it would be morally abhorrent...an embryo or a cloned embryo has a developing and intermediate moral status, certain worthy uses of them may be justified regardless of how and why they were produced. Because the use of stem cells from cloned embryos may in the future provide treatment for serious human diseases, the creation of cloned embryos and their subsequent disaggregation to isolate stem cells can be justified.17

One of the most hotly debated topics in the ethical discourse over embryo research concerns the point in the embryo’s development beyond which embryo experiments must not be permitted. The 1979 report of the Ethics Advisory Board to the Secretary of Health, Education and Welfare claimed that experiments should not be allowed on embryos that had developed beyond the fourteenth day after fertilization.18 This standard was also accepted in 1984 by the Warnock Committee in Great Britain.

One reference point in the development of the human individual is the formation of the primitive streak….This marks the beginning of individual development of the embryo. Taking such a time limit is consonant with the views of those who favor the end of the implantation stage as a limit. We have therefore regarded an earlier date than this as a desirable end-point for research. We accordingly recommend that no live human embryo derived from in vitro fertilisation, whether frozen or unfrozen, may be kept alive, if not transferred to a woman, beyond fourteen days after fertilisation, nor may it be used as a research subject beyond fourteen days after fertilisation.19

This limited status has been enshrined in the Human Fertilisation and Embryology Act of 1990. British embryo researchers cannot receive licenses for experiments on embryos that have developed for more than fourteen days. If they carry out experiments on embryos that have passed the fourteenth day of development, they can face serious criminal penalties.20

Twelve years after the passage of the HFE Act, the pro-research minority on the President’s Council on Bioethics accepted the fourteen day limit as a point beyond which researchers should not pass.

Where to set the boundary is a matter for prudent judgment. For the foreseeable future, the moral line might be safely drawn at fourteen days of development, when no nervous system has developed and when a distinct identity as a single individual has not yet been preordained.21

Not only is there argument over when research is acceptable, but debate also focuses on whether one kind of embryo is usable for research while others might not be. The 1984 report of the Warnock Committee states that embryo research should only be done when necessary, but that if needed, embryos should be created for experiments under strictly controlled conditions. This became the basis of the requirement in the HFE Act that, “No license under this paragraph shall be granted unless the Authority is satisfied that any proposed use of embryos is necessary for the purposes of the research.”22

As the HFEA already existed and had developed legitimacy, many in the House of Lords felt that creating embryos by cell nuclear transfer was no different than by in vitro fertilization, as doing so would be controlled by HFEA licensure procedures. In a speech before the House of Lords on January 22, 2001, Lord Taverne argued that once one decides that creating embryos for research is acceptable, the method of creation is of no significance.

What new moral boundary is being crossed here? At present, new embryos are created by fusing the nucleus from a sperm with an egg. It is now proposed that we should allow the creation of new embryonic cells by transferring a nucleus from another donor cell to the egg. In both cases new embryonic cells are created. To say one is moral and the other immoral is really an argument about angels dancing on the head of a pin.23

This argument did not have the force of law behind it in the United States. However, members of the President’s Council on Bioethics who felt that nuclear transfer experiments should be permitted also depended on it.

...in both cases – creating embryos to aid fertility or creating embryos for biomedical research – the
The fear that embryo research could encourage people to see human life as a mere commodity, rather than something of great value, is a strong theme in the pro-life discourse.

There are other arguments on which the case for embryo science can stand. Virtually every argument made for it in the United States appears in Great Britain and visa versa. The case in opposition to it is equally broad in scope.

The Case against Embryo Science and “Cloning for Biomedical Research”

The argument that it is acceptable to experiment on embryos early in their development, but not later, is highly objectionable to many people in the pro-life movement. For them, the experiments must be seen as either universally acceptable and therefore abhorrent to them or utterly inappropriate. In this section, arguments in opposition to embryo research and nuclear transfer experiments will be reviewed. Many others will not be described either because they are highly technical or directly responsive to the case in support of embryo experiments. However, many arguments are quite general, focusing on the overall status of the human embryo as well as the damage to patients and society in carrying out this research. These will be reviewed in some detail below.

Carrline, et al., in their dissent to the recommendations of the Warnock Committee’s 1984 report claim that any experimentation on embryos is morally unacceptable, wrote:

...The embryo has a special status because of its potential for development to a stage at which everyone would accord it the status of a human person. It is in our view wrong to create something with the potential for becoming a human person and then deliberately to destroy it.

Sir John Peel also writes in opposition to embryo research in an essay published in Embryos and Ethics, a compilation of essays published in 1987. He states that his embryology classes at Oxford were told that the embryo follows a developmental process that moves forward from conception through to birth. Peel, foreshadowing the argument set forth by the majority of the President’s Council on Bioethics who believed that any form of human cloning was unacceptable, claims that the Warnock Commission report looked at the embryo in the reverse. He quotes a statement by Dr. David Woollam, an embryologist, who disagreed with the Warnock Committee’s attempt to determine the last point at which embryo research would be acceptable. Dr. Woollam felt that there was no point at which an embryo was anything less than fully human and he therefore concluded that embryos deserve the same protection as children or adults.

Many Americans also objected to using embryos in research. By the late 1990s, when embryonic stem cell experiments and nuclear transfer research were possible, many focused on the rights of the cloned embryo as a specific issue. In a letter to the New England Journal of Medicine, written in 1998, Professor Robert White of Case Western Reserve University complained that supporters of research cloning did not concern themselves with the value of the nuclear transfer embryos as human life:

...but in your enthusiasm to embrace a new era of biotechnology, you have forgotten that the cloned embryo would be a human being entitled, in spite of its microscopic size, to the same rights and respect granted the fully formed child.

Organs of the Catholic Church were also concerned with attitudes toward human life that could appear as a result of research that caused the destruction of human embryos. In 1999, Richard Doerflinger of the National Council of Catholic Bishops spoke out against supporters of cloning for biomedical research whom he believed were engaging in the commodification of human life.

Catholic moral teaching views human life as a continuum and sees human individuals of every age and condition as meriting the same respect for their fundamental right to life. Differences in mental and physical ability or in stage of development are seen as variations within a larger unity: The human individual, called into existence by God and made in the divine image and likeness, who must always be treated as an end in himself or herself, not merely as a means to other ends.

Writing in 2002, the ten-member pro-life majority on the President’s Council on Bioethics argued that claims that the embryo is deserving of special respect without being viewed as fully human are contradicto-
A year and a half later, the pro-life majority on the President’s Council on Bioethics was also concerned that society could be weakened by “allowing and endorsing the complete instrumentalization of human embryos.” They claimed that cloning for biomedical research would do this because:

In the eyes of those who create IVF embryos to produce a child, every embryo, at the moment of its creation, is a potential child. Even though more eggs are fertilized than will be transferred to a woman, each embryo is brought into being as an end in itself, not simply as a means to other ends. Precisely because one cannot tell which IVF embryo is going to reach the blastocyst stage, implant itself in the uterine wall, and develop into a child, the embryo “wastage” in IVF is more analogous to the embryo wastage in natural sexual intercourse practiced by a couple trying to get pregnant than it is to the creation and use of embryos that requires (without exception) their destruction.

Pro-life activists on both sides of the Atlantic express great concern that embryo research is either being oversold or that the lives of embryos will be sacrificed to corporate greed. In testimony given in March of 1997 before the subcommittee on Public Health and Safety of the Senate Committee on Labor and Human Resources, Professor George Annas of Boston University expressed concern about the risk that cloning might be done for profit:

One does not have to believe in the sanctity of human life, or that life begins at fertilisation, to be concerned about the general commodification of life. Every generation is tempted by the seductive and tantalising prospect of universal happiness as a trump over all other values and principles, but human dignity must always be defended against the abuse of scientific techniques.

Baroness Williams was also concerned by the possibility that the proposed changes in British regulations could mean that human embryos would be seen as tools, rather than the people that many believe them to be: “...using human beings as instruments – what the ethics committee of the European Commission described as the instrumentalisation of human beings.” During her speech, she offers evidence in support of her concerns by pointing to changes in the kinds of people who offer organs for donation and the ethical concerns that have arisen as a result:

If we have any doubts about that danger, we need only consider the fact that human organs are being sold on the Internet for use as spare parts. Those who can afford it buy them for large sums of money from their unwilling donors, who are too poor to be able to finance their lives in any other way.
I use that word deliberately—insatiable public demand. If the doctors say it, it must be done. Are we to permit research to begin without considering the social use to which it will be put if it proves successful? Who is to own the patent? Are we to devote the product of human beings to the control of a large pharmaceutical company?40

Objections from opponents to embryo research like Professor Annas and Lord Brennan were often paired with expressions of concern about the possible harm that could come from promising results to patients where much research is yet to be done. In her speech before the House of Lords in January of 2001, Baroness Blatch quotes a letter from a group of disabled people in Britain who object to using their conditions as a tool for gaining political support for embryo research.

We object strongly to the propaganda that is being used to claim that virtually our only hope of cures lies in the use of human embryonic stem cells. This, of course, is not only totally untrue, but it is an extremely cruel method of gaining support. It amounts to blackmail—exploitation of the very worst kind—of disabled and sick people, some of whom are in extremely difficult situations.41

A year and a half later, the pro-life majority of the President’s Council on Bioethics were equally forceful in expressing concern about the possible harm that could come from promising results to patients where much research is yet to be done. They claimed that nuclear transfer to create embryonic stem cells might fail to bring results while less morally questionable research, such as with adult stem cells, might lead to great advances without the need for the destruction of human (cloned) embryos.42

Three years after the publication of the report of the Warnock Committee in 1984, British ethicists expressed concern about the possibility that researchers and clinicians might devalue children who had genetic defects. George Chalmers, writing in a paper published in Embryos and Ethics, complained as early as 1987 that, “prevention of genetic defects—genetic engineering” was part of a “catalogue of arrogant interference” whose scope he believed was infinite should embryo research be allowed in Great Britain.43

Fifteen years later, the pro-life majority on the President’s Council on Bioethics expressed great concern that cloning for biomedical research could weaken society’s underpinnings and devalue the norms that it hands on to future generations. They were concerned that society would be weakened, “by crossing the boundary from sexual to asexual reproduction, in the process approving, whether recognized or not, genetic manipulation and control of nascent human life.”44

Some opponents of various kinds of research fear that once one kind of experimentation is considered acceptable, others will then be sought by researchers who see value in carrying them out. This argument, often called the “slippery slope,” frequently appears on both sides of the Atlantic.

In their expression of dissent to the Warnock Committee’s 1984 report, Carriline and others argued that this was a risk even with procedures that were allowed by law:

…it is necessary also to look beyond the procedure to possible consequences. The proposal that the creation of embryos for research should be banned but that experiments on embryos created to relieve infertility which are no longer required for that purpose so called “spare” embryos be permitted, comes into this category. In our view experiments on “spare embryos” are wrong. But even if they were deemed right, the consequences of permitting them would be unacceptable. There would be a strong temptation for doctors to harvest more embryos than strictly required for the immediate therapeutic purpose in order to provide “spare embryos.” “Spare” would become a euphemism.45

Seventeen years after Carriline, et al., expressed their fears on this matter, Lord Alton claimed that the Human Fertilisation and Embryology Act had proven to be a beginning and the newly proposed extensions to its remit to allow study to cure “serious disease” showed that as technology advanced the demand to do more had increased. In his January 22, 2001 speech in the House of Lords, Lord Alton says:

...even at this early stage of development, we are not dealing with something that is inconsequential. There is nothing therapeutic in this procedure for the new human embryo: once it has been used, it will be destroyed. There is no question here of donor rights or consent. Since 1990,...between 300,000 and half a million human embryos have been destroyed or experimented upon....our willingness to walk this road has paved the way for more and more demands.46

The ten-member pro-life majority of the President’s Council on Bioethics also had fears that one set of possibilities that might be acceptable now might lead to
other, less acceptable possibilities later. It laid its argument that societal health could be damaged by allowing any human cloning at all directly on its fear that society could begin a trip down a slippery slope....opening the door to other – for some of us, far greater – moral hazards, such as cloning-to-produce-children or research on later-stage human embryos and fetuses.47

Along with the risk that new procedures might be demanded comes the risk that old methods may be used in new and socially unacceptable ways. Sir John Peel, writing in the 1987 compiled volume Embryos and Ethics, expressed deep concern for the health of British society and the implications that freely available in vitro fertilization might have on it:

These techniques can be and are being extended outside marriage – even to single women with no male partner, who are anxious to have a child, but who wish to avoid involvement with any member of the opposite sex, moral, emotional or legal. Are the wishes of the individual woman always to be paramount? It is not for doctors to provide answers to such a question, but I believe society should give far more serious consideration to the direction in which we are moving and consider whether it is not laying the foundations of problems to come.48

Fifteen years later, the pro-life majority of the President's Council on Bioethics felt that cloning for any purpose could not be allowed because it risked far too much damage to the values that they believed should be of fundamental importance to society.

In trying to discern where a wise and prudent boundary must be drawn – to protect those beings who are humanly inviolable, to prevent the dangers that most tempt us, and to protect the moral fabric of society – we hold that the boundary must be drawn by prohibiting the production and use of cloned embryos. To cross this boundary or to set it further down the road – that is, “with limits” – is to invite (and perhaps ensure) that some (or all) of the dehumanizing possibilities described above will come to pass.49

Unwilling participation in the destruction of human embryos due to government regulations that require this action is also highly objectionable to members of the pro-life movement. Writing in his 1997 book, The Embryo Research Debate, Michael Mulkay quotes a piece by Peter Garrett in the Summer 1995 issue of Life News that expresses the anger that many in Britain had over the mandated destruction of embryos as required by the HFE Act:

The 14-day time-limit for human embryos is in effect an order to execute. The law dictates that the passing of the fourteenth day confers ‘humanity’ on the embryo and requires that it be destroyed. The appearance of the primitive streak, so beloved of Anne McLaren and Mary Warnock, is both the moment when, allegedly, humanity appears and when the new human being must be destroyed.50

American pro-life activists were equally concerned by the possibility that research cloning might be allowed, while reproductive cloning was banned because of the risk that society would mandate the destruction of nuclear transfer embryos that it found out were to be used for reproductive purposes. In a June 6, 2002 letter to Senators expressing its opposition to “clone and kill” legislation, the National Right to Life Committee says:

enactment of a clone-and-kill bill such as S. 2439 would allow and in fact encourage all of those same evils [creation of embryos for their ultimate destruction and use as model organisms] and then add one new and highly objectionable component:...It would place the FBI and other federal law enforcement agencies in charge of keeping track of countless cloned human embryos and ensuring that none of them survive. Indeed, under the forfeiture provision of S. 2439, in some circumstances federal agents could be forced to directly seize and destroy cloned human embryos (“the product of nuclear transplantation”), to enforce the bill’s ban on human clonal pregnancy and birth.51

Conclusion

Embryo research is a particularly powerful lightning rod for debates over how and whether scientific knowledge should be used because it can help people to have children, and may help patients who would be sick or die without the knowledge that can be gained from it. However, societal values can be threatened, and the tension between these tremendously powerful forces will continue to force policy makers to make decisions that will make many uncomfortable. The most important lesson that can be taken from this review of the ethical arguments made in support of and opposition to embryo research in the United States and United Kingdom is that different countries will make different decisions when faced with the same ethical choices. As this is the case, it seems important to begin asking not whether knowledge
should be used, but rather what society will do in response to its existence and the likelihood that someone will try to apply it to human health. Only when we begin framing our choices in this way will we be asking questions that mirror those that clinicians, researchers, patients and their loved ones ask themselves before deciding that they want to access a new capability that offers them hope no matter the objections that others may have to their choices. These choices will have a profound impact on biomedicine and other applications of biotechnology in the future because nation states that hope to prevent or radically alter the growth of biomedical knowledge will have increasing difficulty controlling their citizens tightly enough to prevent many of them from exploring tools and techniques that make many uncomfortable. We are fortunate in that technologies that do not work will ultimately not be used, and it is possible that we should trust its consumers – ourselves – to make the decisions that will ultimately serve all of us best, as making them for other people will become increasingly difficult as time passes.

Acknowledgements
Research for this article was supported by the National Institute of General Medical Sciences under NRSA no. 1 F31-GM20933.

References
4. PCB, supra note 3.
6. HFEA, supra note 1.
7. Id.
8. Warnock, supra note 5, at 64.
11. Id.
14. PCB, supra note 3.
15. Patel, supra note 13, column 58.
17. PCB, supra note 3.
19. M. Warnock, supra note 5, at 64.
21. PCB, supra note 3.
22. HFE Act, supra note 20.
24. PCB, supra note 3.
27. Id.
29. Id.
32. PCB, supra note 3.
35. Id.
36. PCB, supra note 3.
37. Id.
42. PCB, supra note 3.
44. PCB, supra note 3.
47. PCB, supra note 3.
48. Peel, supra note 28.
49. PCB, supra note 3.
Several high-profile episodes have recently thrust drug safety and the pharmaceutical industry’s practices into the spotlight. Merck’s recall of the drug Vioxx, for instance, was a major news event. GlaxoSmithKline’s suppression of data linking suicidal behavior among children to Paxil also galvanized tremendous public attention. What differentiates these events from the usual evolving process of scientific knowledge, and marks them with an aura of “scandal,” are questions about the propriety of corporate behavior. Who knew what, and when did they know it? Concerns are growing about the potential for industry sponsors to suppress negative results from clinical trials research. Scientists, medical journal editors, legislators, and the public have called for greater transparency in the conduct of clinical trials and the drug approval process.

One reform that has gained increasing support is a clinical trials registry, designed to act as a comprehensive repository for information on ongoing and completed trials. In the spring of 2004, the International Committee of Medical Journal Editors (ICMJE) signaled its strong support for this reform by announcing that registration in such a repository will become a condition for consideration of any manuscript that reports results of a trial. The explicit goal is transparency in the clinical trials process.

The ICMJE requirement is set to apply all trials commencing enrollment after July 1, 2005. It further stipulates that the registry used must be managed by a non-profit organization. In practice, this means that investigators who wish to publish their results from a clinical trial in a high-impact journal will need to register details of the trial such as hypotheses, outcome measures, and eligibility criteria, at the trial’s inception, before patient enrollment begins. Registration would create a public record of the trial’s existence that would be widely accessible. Although the ICMJE comprises only eleven member journals, editors from the leading general medical journals are represented, and standards set by this body carry considerable influence over scholarly publications in the clinical sciences.

In this paper, we support the establishment of a prospective registry for clinical trials as a means of promoting drug safety and scientific integrity. First, we outline why clinical trial registration is important today. Second, we describe registry initiatives by government and private organizations. Third, we outline several important limitations to existing registries. Finally, we propose a set of elements that we believe are integral to a successful clinical trials registry.

Why Register Clinical Trials?

The Importance of Clinical Trials

The fundamental principle underlying the need for the registration of clinical trials is that this particular research mechanism has special scientific and public importance. The ICMJE has defined a clinical trial as “any research project that prospectively assigns human subjects to intervention or comparison groups to study the cause-and-effect relationship between a medical intervention and a health outcome.” Randomized controlled trials (RCTs), a subset of clinical trials,
are especially important for demonstrating relationships between interventions and outcomes. Since the first large randomized trials of streptomycin and antimalarial agents in the 1940s, the RCT has been regarded as the “gold standard” for testing the efficacy of medical interventions.

Countless advances in medicine owe their existence to the scientific rigor demonstrated by comparing clinical outcomes in “cases” against “controls,” with the randomized selection of both populations largely eliminating the possibility of alternative causality. As a recent example, Dickersin and Rennie point to the results of the Health and Estrogen/Progestin Replacement Study and the Women’s Health Initiative Study as evidence for the superiority of the RCT design.

Those studies both provided startling evidence that hormone replacement therapy was in fact not effective in reducing the risk of cardiovascular disease in postmenopausal women, and may even have been harmful. Such findings went against years of evidence gleaned from observational studies, demonstrating that even in the face of overwhelming observational support for a given intervention, a RCT is the only way to conclusively demonstrate efficacy. Several commentators have recently argued that until rigorous RCT methodologies are extended to organization-and service-level interventions in health care, the true efficacy of these measures cannot be known.

Unpublished Trials and Publication Bias

Despite the primacy of RCT methods, and the critically important nature of their outcomes, mechanisms for systematically disseminating their results – indeed, even indicating of their very existence – are limited. Studies have indicated that about half of presentations at scientific conferences never translate into published articles, and these numbers include clinical trials. The peer-review process may reject some of these reports on the basis of poor research design or lack of scientific merit. However, the explanation for non-publication in vast majority of cases appears to lie with the investigator; many are simply never written up or submitted. Some commentators have characterized this behavior as a form of scientific misconduct.

Failure to publish clinical trial reports is problematic because there is evidence that the results of unpublished trials are systematically different from their published counterparts. Specifically, findings that show efficacy of the intervention under study are more likely to be published, and vice versa. A clear correlation between non-publication and lack of significant findings has been demonstrated. Again, journal editors may have a hand in this result, but investigator unwillingness to submit such reports for publication appears to be key factor. Whatever the precise cause, the result is clear: published studies frequently favor the new drug or intervention under study, and reports of clinically promising interventions are overrepresented in the medical literature. The phenomenon is commonly referred to as “positive publication bias.”

Positive publication bias has several unfortunate consequences. It is inconsistent with the push toward evidence-based medicine. The evidence base for clinical practice and physician decision-making is weakened considerably by the absence of relevant information. Similarly, systematic reviews of the literature and meta-analyses regarding a given intervention will be skewed. Any given intervention may have undergone multiple studies, some showing benefit and others showing no benefit. The goals of systematic reviews and meta-analyses are to weigh the opposing evidence, separate relevant from irrelevant findings, evaluate the quality of each study, and arrive at a judgment about efficacy in light of the available evidence. The absence of studies showing no effect may tip conclusions inappropriately toward a judgment of efficacy and, indirectly, stimulate utilization patterns that adversely impact quality of care.

At the policy level, positive publication bias threatens to deprive regulators and policymakers of the information needed to make informed choices. Decisions ranging from drug formularies to approval of new devices to screening recommendations may be affected. Consumers stand to lose in such circumstances. Thus, publication bias jeopardizes sound decision making at multiple levels within the health care system – from clinical judgments at the bedside to policy decisions in legislatures or consumer protection agencies, and many places in between. A well-functioning clinical trials registry would guard against this threat by providing an authoritative and
complete account of ongoing and completed trials. Importantly, both positive and negative findings would be readily accessible.

**Suppression of Data**

A more malignant form of failure to publish negative results is data suppression. In this situation, investigators’ failure to submit their research results for review and publication is driven not by apathy but by an active, considered, and intentional decision not to disseminate them. Such behavior clearly violates professional ethical norms in the biomedical and clinical sciences; it may also constitute fraud.

The recent case of *Spitzer v. GlaxoSmithKline PLC* is an example of trial reporting gone awry. On June 2, 2004, New York Attorney General Elliot Spitzer filed suit against pharmaceutical giant GlaxoSmithKline (GSK), manufacturer of the popular antidepressant Paxil (paroxetine). In the year prior to the commencement of legal action, there had been considerable controversy in the United States and United Kingdom over whether children with major depressive disorders had an increased risk of suicide when given a class of antidepressant medications known as Selective Serotonin Reuptake Inhibitors (“SSRIs”); there were also lingering questions about whether these drugs were in fact more effective than placebo in the treatment of pediatric depression. Although the Food and Drug Administration in the US had not officially approved Paxil for use in children, doctors could prescribe the drug to children through “off-label” use, and they did: approximately 2.1 million Paxil prescriptions were written for children nationwide in 2002.

During the course of the litigation, evidence emerged implicating GSK in suppression of data about Paxil’s efficacy. An internal company memorandum written in 1998 noted that two of three clinical trials conducted in pediatric populations had failed to demonstrate any differences between Paxil and placebo. However, with respect to company policy and existing research on Paxil, the memorandum stated that “it would be commercially unacceptable to include a statement that efficacy had not been demonstrated, as this would undermine the profile of paroxetine.”

Spitzer’s allegations also pointed out that in June of 2003, the UK’s Medicines and Healthcare products Regulatory Agency had announced that its analyses of GSK’s data suggested that the risk of potential suicidal behavior in children with depression was significantly higher in the Paxil group than the placebo group. GSK issued a press release in Britain publicizing this finding, but an American press release of the same month stated that “there is no evidence that Paxil is associated with an increased risk of suicidal thinking or acts in adults.” The release was silent on the subject of ill effects in children.

Thus, Spitzer’s case against GSK alleged that information disseminated on Paxil in the US was deficient in failing to acknowledge studies that had raised significant questions about both the drug’s efficacy and its safety in pediatric populations. Spitzer argued that the company had a duty to disclose negative studies, not just positive ones. GSK settled the case on August 26, 2004. The settlement terms excluded an admission of wrongdoing, though GSK committed to submitting all future clinical trial results to a clinical trials register to be posted on the Internet.

The Paxil case illustrates the tension that may arise between private and public interests in access to information about the safety and efficacy of clinical interventions. Although commercial entities such as drug and device manufacturers are in the business of improving healthcare, their corporate goals center on financial success. When the two objectives conflict the potential for suppression of data arises, with possible adverse consequences for consumers. This scenario presents the classic argument for regulatory intervention in the marketplace. In this case, mandatory registration of clinical trials would help to ensure that the wider public interest in information from clinical trials was not subordinated to private interests in sale, profit, and commercial reputation.

**Disaggregation and “Salami Science”**

The twin problems of positive publication bias and data suppression lead to a deficiency of information. A reverse problem — namely, oversupply of information — also poses a threat in certain circumstances. In 1986, Edward Huth referred to the practice of publishing multiple papers on a single study as “salami science”; in other words, “slicing” reports of a single trial to give the illusion of greater support for the intervention under study. This phenomenon, more commonly known as “disaggregation,” occurs when results of what is essentially a single trial are presented or interpreted as discrete studies in multiple publications. Reports arising from multicenter trials are especially prone the illusion. Interest in disaggregation may be fuelled both by those who wish to gain from the appearance of a large body of support for a given intervention and academics for whom publication volume stands as key indicator of productivity.

An investigation by Patricia Huston and David Moher’s into studies of the psychiatric drug risperidone provides an interesting case study of disaggregation of RCT results. Two major and seven minor trials...
had been conducted on risperidone at the time of their investigation. A literature search revealed 20 articles and several unpublished reports describing RCTs of the drug. Huston and Moher found "obvious redundancy" in the results of a single center study; despite most biomedical journals' prohibitions against duplication, the findings had essentially been published twice in different journals. Detecting disaggregation in the multicentre trials was much more challenging task due to "the chronology of publications, changing authorship, lack of transparency in reporting, and the frequent citation of abstracts and unpublished reports."

Nevertheless, Huston and Moher found considerable disaggregation in a multicentre study of risperidone conducted in Canada and the US. In 1992, the results of the entire North American study were first published as an abstract in Clinical Neuropharmacology. In 1993, the same study was published in Schizophrenia Research while the Canadian results were separated and published in the Journal of Clinical Psychopharmacology. No mention was made of any affiliation with the larger North American Trial, making it unclear as to whether this study was in fact part of the larger North American one. However, in 1994, the American results alone were published in the American Journal of Psychiatry; the investigators did mention that their study was part of a larger North American trial, and cited the publication of the Canadian results. Also in 1994, a minor reanalysis of the North American results was published in the Journal of Clinical Psychiatry. Finally, Huston and Moher identified several references to an unpublished report of the full North American trial results. Hence, the results of a single multicentre study were published a total of six different times, with varying authorship for each paper.

Disaggregation has several serious implications. First, it confounds editors and the peer review system upon which the publication process for original research is founded. Second, it may artificially inflate the apparent evidence for given intervention. The risperidone example is instructive in this regard: following uncritically the pattern of the published reports, advertisements for the drug risperidone in the US referred to North American and Canadian trials. Finally, disaggregation may exacerbate the problem of publication bias. Any misimpression of support for interventions that is created by the absence of opposing studies will be compounded by an artificially large number of supporting reports.

A clinical trials registry could attenuate the problem of disaggregation by providing a clear enumeration of the trials conducted on a specific intervention. Publication of overlapping reports, or interpretation of such reports as discrete findings, would be more difficult if editors, peer reviewers, and regulators had an easy point of reference to check the study's history and publication record. Determining where one trial ends and another begins will not always be straightforward. Institutional review boards confront this challenge frequently. Relevant factors in determining discreteness include the research objectives, identity of the study sample, nature of the intervention, location of the research, and funding sources. A registry would arm those attempting to interpret the novelty and value of trial results with a powerful tool.

**Public Accessibility & Accountability**

Members of the general public are important stakeholders in clinical trials. This interest is self-evident in the case of publicly-funded trials. However, it also extends to privately-funded trials, which turn to patients for the enrollment that makes the research possible. Moreover, private funders inevitably rely on consumer patronage to make commercial successes of treatments that have demonstrated efficacy in clinical trials.

Registration of clinical trials honors the public’s stake in their conduct and outcomes. Registries provide a mechanism through which members of the public can learn of trials in which they, family members, or friends may wish to participate.
groups were the driving force behind the establishment in 2000 of the website clinicaltrials.gov, a clinical trials registry containing information on trials pertaining to serious or life threatening conditions.

Registries also serve an accountability function, linking the medical research community to the public they serve. In this sense, the research may be construed as a type of social contract. Clinical advances depend on the voluntary participation and goodwill of research subjects. In return, the scientists and sponsors involved bear an ethical (if not legal) obligation to the individuals who make the work possible. The fact that research subjects often cite notions of the “greater good” as an important motivation for participation reinforces the notion that reciprocal responsibilities underpin the clinical research enterprise. Some reasonable level of transparency, of the kind that a comprehensive registry could provide, would seem to be a fundamental step in discharging those responsibilities.

Existing Clinical Trials Registries
A number of publicly-accessible registries of clinical trials currently exist. In this section, we describe several of the leading ones.

ClinicalTrials.gov
ClinicalTrials.gov is a web-based registry administered by the National Library of Medicine. Mandated by a section in the Food and Drug Administration Modernization Act of 1997, the registry was created in response to demands by patient advocacy groups for greater access to clinical trials. The Act called for a "registry of clinical trials...of experimental treatments for serious or life threatening diseases...in a form that can be readily understood by members of the public." The site became active in February of 2000. Although the Act stipulates that registration of all clinical trials relevant to serious or life threatening diseases is mandatory, no enforcement mechanisms are set forth.

Entries on ClinicalTrials.gov provide a number of pieces of information about each trial, including: a summary of the purpose of the study; recruiting status; eligibility criteria; trial location; study design; trial phase; disease or condition and drug or therapy under study; and links to health resources, such as MEDLINEplus and PubMed, that “help place clinical trials in the context of patients' overall medical care.”

Currently, ClinicalTrials.gov contains approximately 11,900 clinical studies sponsored by the NIH, other federal agencies, and the private sector. Studies listed in the database are conducted in all 50 states and in over 90 countries. ClinicalTrials.gov receives over 3 million hits per month, and approximately 17,000 visitors daily.

metaRegister of Controlled Trials
The metaRegister of Controlled Trials (mRCT) is a free, searchable online database of RCTs. Based in the UK, and administered by publisher BioMed Central, it is made up of registers held by public, charitable, and commercial sponsors of trials. Currently participating registers come from the UK, Canada, the US, Australia and Hong Kong.

According to its website, mRCT is intended for: individuals who use evidence from RCTs and want to ensure a comprehensive knowledge of all evidence on a subject; funding agencies who wish to guard against duplication in research and funding, or make funding decisions in light of all evidence; and members of the public looking for information about trials in which they may wish to enroll. The International Standard Randomized Controlled Trial Number (ISRCTN) registry is also administered through this registry. An ISRCTN is a unique number that allows a trial to be tracked and identified through out its life cycle; it is a way to keep track of trials internationally. In addition to the ISRCTN, mRCT lists the following information about each registered trial: title; sponsor; disease or condition under study; hypothesis and objectives; eligibility; current status of trial; and contact information.

Industry Sponsored Registries
In the wake of the GlaxoSmithKline lawsuit, the Executive Committee of the Pharmaceutical Research and Manufacturers of America (PhRMA), the trade association that represents the major pharmaceutical manufacturers, adopted a set of principles regarding transparency, communication, and clinical trial results. To facilitate “timely communication of study results, regardless of the outcome of the study,” PhRMA established an electronic database of study results, accessible at www.clinicalstudyresults.org.

One company, Eli Lilly, independently and established its own more comprehensive registry. Unlike the PhRMA registry, Lilly’s website includes not only the results of completed studies, but also a description of those studies which are ongoing or recruiting. All Lilly-sponsored Phase II, III, and IV clinical trials initiated after July 1, 2004 are listed. The company states that, “Posting trials in this initiated trials section is a commitment by Lilly to ensure that all trials that are started will have disclosure of the results, regardless of outcome.” The inclusion of all initiated trials in the Lilly registry makes it more useful than the PhRMA registry.
On January 7, 2005, four large pharmaceutical industry trade groups reiterated a commitment to releasing more information on clinical trials, including completed trials. PhRMA was among them, as were similar organizations from Europe and Japan. The commitment pledged that any member company commencing a clinical trial would list specified information about it on ClinicalTrials.gov. Companies would be encouraged to release results of a trial within a year of a drug’s approval, and the results would appear in a publicly accessible database run either by the government or a company or trade group. The reach of the proposal, however, remains modest in several key respects. Certain types of data, such as information on Phase 1 trials, are exempted. More importantly, the effort remains entirely voluntary.

Limitations of Existing Registries
Although the establishment of registries to date by public and private entities signals a commitment to the wider dissemination of information about clinical trials, existing registries are fraught with problems. For example, ClinicalTrials.gov covers only trials that pertain to serious or life threatening conditions; trials testing interventions to alleviate lesser conditions are not listed. Further, the section of the Food and Drug Administration (FDA) Modernization Act governing ClinicalTrials.gov does not provide for any enforcement mechanism should an investigator fail to register a trial. This makes it difficult to secure compliance, and holes in compliance are evident. Of 11,900 trials currently registered on ClinicalTrials.gov, only 2230 (less than 20%) are sponsored at least in part by industry. Given that greater than 60% of all clinical trials are currently funded by industry, this is a low number. Indeed, an FDA review demonstrated that between January and September 2002, among trials testing cancer treatments, 91% of government-sponsored trials and 49% of industry-sponsored trials had been registered.

In October of 2004, Democrats introduced into both houses of Congress a bill designed to address shortcomings with the reach and operation of ClinicalTrials.gov. The Fair Access to Clinical Trials (FACT) Act proposes mandatory registration of all clinical trials conducted in the United States. It would require the reporting of such details as research outcomes, basic demographic information, sources of funding, significant adverse events, and FDA approval status. Most importantly, The FACT Act gives oversight of this clinical trials registry to the Secretary of Health and Human Services. Strong enforcement measures are provided for, including financial penalties of up to $10,000 per day for refusal to comply. In addition, the FACT Act mandates participation in the registry as a prerequisite for approval by the local Institutional Review Board (IRB). Because IRB approval is a legal requirement for virtually all research activities conducted in the U.S. that involve human or animal subjects, these bodies are ideally placed to act gatekeepers and help to ensure widespread adherence to the registration mandate.

On December 10, 2004, Charles E. Grassley, a Republican senator, announced his intention to introduce legislation in 2005 requiring drug companies to register clinical trials and report results in a public database. He indicated that the forthcoming bill is likely to echo some of the features of the FACT Act. The involvement of Senator Grassley is significant because it suggests growing bipartisan support for sterner legislation in this area.

The mRCT registry has some distinct advantages over ClinicalTrials.gov. Its international scope is especially noteworthy. However, it also has several notable shortcomings. Although the database is freely accessible to the public, it is privately owned by BioMed Central, a for-profit company. A lively debate about the acceptability of for-profit registry ownership has played out over the past year. The ICMJE statement on clinical trial registration includes public or non-profit ownership as a requirement of an acceptable trials registry.

With respect to the issuing of ISRCTNs, also administered through mRCT, registering with an ISRCTN was recently mandated by the World Health Organization (WHO) for all WHO-approved RCTs. To date, ISRCTNs have been assigned to over 1800 trials. It is problematic, therefore, that there is discord between the requirements of the ICMJE and mRCT/ISRCTN. The British Medical Journal has singled this out as a reason why it has not signed on to the ICMJE requirements, instead choosing to adopt their own set of criteria for a clinical trials registry. The acting editor, Kamran Abbasi, states that while public ownership is valuable, free public access should suffice.

A further concern regarding the ISRCTN is that it costs approximately $144 to register unless trials are “sponsored, funded and carried out in developing and transitional countries.” While this is far from prohibitively expensive, in the context of a system that relies on voluntary action, even a nominal fee may undermine the goal of universal registration by discouraging some prospective registrants.

With respect to industry-run registries, several concerns are identifiable. First, the PhRMA registry is voluntary, with no consequences for failing to register. This will likely result in a low registration rate. Even if a company is not opposed to posting the results of a
Given clinical trial, the fact that it is not mandatory means that it might end up at the bottom of a to-do list. Second, the registry is not prospective, which may undermine key goals of registration. As the current goal of establishing trials registries is to ensure negative trials are not hidden *ex post facto*. Finally, the logic of having industry run a clinical trials registry is questionable. Says *JAMA* editor Catherine DeAngelis: “Why would you put the fox in charge of the henhouse? And, if they have nothing to hide, why waste money setting up their own database?”

Another problem associated with industry-run registries, particularly retrospective ones, is the potential for selective registration. Even if registration rates are very high, there must be ongoing concern that the relatively small number of trials *not* registered are especially sensitive.

### Key Elements of an Optimal Clinical Trials Registry

For the reasons outlined above, we believe that none of the existing clinical trials registries fuses the ideal set of design features. To ensure that registration is complete, meaningful, and useful to the various stakeholders, we propose five core design features. A registry should be: (1) international in reach; (2) prospective in nature; (3) owned and operated by a public or non-profit entity; (4) part of a coordinated, systemic effort between interested parties; and (5) mandatory. In the remainder of this section, we elaborate each of these key elements of an optimal clinical trials registry.

### International

In an era of globalization, trials are increasingly multicenter in design and cross geopolitical boundaries. For example, many trials of HIV/AIDS therapies are conducted in Asia and Africa, where pathophysiological attributes of the virus differ and its scourge is felt most acutely. A single, worldwide clinical trials registry (or meta-registry) would ensure the reporting and accessibility of all trials, regardless of location. Such a registry could be administered by a public or non-profit organization experienced in the coordination of international research information such as the Cochrane Collaboration. The Cochrane Collaboration is an international, non-profit group that “produces and disseminates systematic reviews of healthcare interventions and promotes the search for evidence in the form of clinical trials and other studies of interventions.”

An international registry would consolidate information, making it easier for scientists and the public to access the information they need in a way that is quick and uncomplicated.

The major challenge facing an international registry is the logistical problem of global enforcement. The coordinating body would not have authority or jurisdiction to legally mandate registration. One possible solution to this problem would be to enshrine registry requirements in international law; participating countries would formally establish their obligations through a treaty instrument.

An alternative approach to the enforcement problem would be to encourage nations in which trials are being conducted to enact domestic laws mandating registration. A third possibility would be to have countries mandate their own stand-alone registries through domestic law, and establish a meta-registry to gather information from the national registries. Hence, a country would operate its own domestic registry, incorporating whatever preferences it had regarding registration and making registration mandatory.

A core set of fields from this registry would be then periodically transferred for inclusion in the international registry.

The obstacles posed by coordinating registry requirements across national lines loom as quite formidable. From a practical standpoint, however, conformity of relatively few nations – those that host or fund that vast majority of clinical research – would secure significant global cooperation. If the United States, the United Kingdom, France, Germany, and Canada were to endorse strongly mandatory reporting of a minimum dataset to an international meta-registry and lead by example, the lion’s share of clinical trials underway today would be covered because they are either conducted in these countries or financed by organizations headquartered there.
Moreover, if these countries enthusiastically endorsed adherence and lead by example, the international pressure to conform would be considerable.

**Prospective**

Unless trials are registered at their inception, the goals of raising public awareness and reducing publication bias will be difficult to achieve. The only sure way to guarantee that the existence of all trials is acknowledged and documented is to mandate prospective registration. Retrospective registration does not aid patients interested in becoming enrollees. It also creates conditions ripe for the conflicts of interest, where those in possession of findings that run contrary to their interests have incentives to bury the findings, whether through active or passive means. Early registration and prospective tracking of trials’ progress is insurance against such behavior which benefits all stakeholders, including, arguably, the very parties who may otherwise be tempted to withhold information at a later point in time. A comprehensive registry with prospective information also provides a means of guarding against the effects of positive publication bias.

**Non-Profit**

The third component of a successful trials registry is non-profit ownership and administration. We believe that a registry should be owned, sponsored, or managed by a neutral third parties, not private groups with a vested interest in the registry’s content. For this reason, registries owned and operated by the pharmaceutical industry are inherently problematic. Though they may be useful supplements to a broad public registry, they are not viable substitutes.

The strictures that apply to the identity of the “neutral” third party is a more nuanced question. There has been debate about whether a for-profit company that guaranteed free access could legitimately own and administer a trials registry. The freely-accessible mRCT, as noted above, is owned and operated by the for-profit publishing group BioMed Central. The editor of the *British Medical Journal* has publicly endorsed this arrangement, arguing that a registry need not be non-profit or public, so long as it can be freely accessed by the scientific community and the public.

We take a less sanguine view for two reasons. First, a for-profit operator/administrator will face tremendous difficulties in avoiding conflicts of interest. For the registry to operate as a viable for-profit entity, it requires a stable source of revenue. User fees would inhibit access. Complete public funding is highly unlikely, and would obviate the value of a non-governmental operator. So where will revenue come from? The mRCT registry provides a interesting case study. On a recent visit to its website, an advertisement for “Adsumo” flashed at the top of the screen. Adsumo is a job-seeking page for those in the biomedical sciences. The featured employers included AstraZeneca, Wyeth, and Amgen. Links to drug companies through advertising revenue diminishes the appearance of neutrality and undermines the goals of a clinical trials registry.

This leads directly to the second argument against for-profit involvement in the management of the registry. Even if diligent and entrepreneurial efforts by a for-profit manage to avoid compromising sources of revenue, it must be remembered that the potential for secondary interests to overwhelm primary interests are the essential ingredients of a conflict of interest. Put another way, the game is one of perception, and the reality of for-profit operation is that skepticism would surround the for-profit operator’s every move. To avoid this untenable situation, a successful clinical trials registry should be owned and operated by a public or non-profit agency.

**Coordinated**

Any clinical trials registry must be the product of a coordinated effort between interested parties for it to function successfully. For example, as discussed above, the ICMJE member journals will soon refuse to consider trials for publication that have not been registered with a clinical trials registry. However, the FDA does not condition drug approval on either registration or publication of trial results in a medical journal. Therefore, the absence of FDA standards that complement those of the ICMJE fragments the regulatory framework. Cooperation among regulatory and professional agencies is needed to avoid the loopholes that often plague parallel systems of governance.

**Mandatory**

Finally, a clinical trials registry must be mandatory in order for it to operate successfully. In other words, registration must be required by legislation comparable to the FACT Act. Opponents of mandatory registration argue that it may lead to revelation of manufacturers’ proprietary information, or impede drug development by forcing the release of research strategies to competitors. Proponents counter that those risks must be weighed against the benefits of registration, and that registration can be meaningful only if it is comprehensive. This camp doubts that voluntary efforts can work.

The pessimism about voluntary registration has some basis in experience. In May of 2003, the
Association of the British Pharmaceutical Industry announced it had established a voluntary register of retrospective Phase III and ongoing Phase IV clinical trials in the United Kingdom.50 Fifteen months later, information about only 93 trials involving 44 drugs tested during the period of 1991-2004 had been posted. Further, as previously noted, less than 20% of those trials currently registered on ClinicalTrials.gov are sponsored at least in part by industry. Between January and September 2004, non-commercial sponsored trials registered each month with ClinicalTrials.gov outnumbered those registered by commercial sponsors until July – right around the time that Elliot Spitzer initiated litigation against GlaxoSmithKline. This suggests that without the threat of litigation or some other “push,” industry will be reluctant to voluntarily register all trial data.

Conclusions

Drumond Rennie has stated that the idea of clinical trials registration has gone from ignored to “irresistible.” We agree. Recent events highlight the conflicted role of industry sponsors in clinical trials research, and emphasize the need for greater transparency and accountability. The significant potential for harm and abuse in this area underscore the necessity of establishing a registry for all clinical trials. The establishment of a registry according to the criteria proposed here – namely, one that is international, prospective, non-profit, coordinated, and mandatory – would promote evidence-based medicine, scientific integrity, public accountability, and the timely access to information on clinical trials worldwide. It is a reform that is past due.

References

2. Ibid.
3. Ibid.
4. In a RCT, subjects are randomly allocated to either the treatment group, or the control group; the control group receives a placebo in lieu of the actual intervention. RCTs may be open (i.e., both the researcher and patient know which group the patient is in), single-blind (the researcher knows but the patient does not), or double-blind (neither party is aware).
12. See Dickersin and Rennie, supra note 6.
16. Systematic reviews and meta-analyses compile the results of all clinical trials on a particular drug or other intervention, and re-evaluate its effectiveness using a greater amount of data.
18. Ibid.
19. Ibid.
20. Ibid.
25. Ibid.
27. Section 113, Information Program on Clinical Trials for Serious Life-Threatening Diseases, Food and Drug Administration Modernization Act of 1997, Public Law 105-115.
29. Ibid.
30. Ibid.
35. Phase I studies are designed to establish the effects of a new drug in humans. These studies are usually conducted on small populations of healthy humans to specifically determine a drug’s toxicity, absorption, distribution and metabolism. After the successful completion of phase I trials, a drug is then tested for safety and efficacy in a slightly larger population of individuals who are afflicted with the disease or condition for which the drug was developed. This is known as a Phase II study. The
third and last pre-approval round of testing of a drug is conducted on large populations of afflicted patients. Phase III studies usually test the new drug in comparison with the standard therapy currently being used for the disease in question. The results of these trials usually provide the information that is included in the package insert and labeling. After a drug has been approved by the FDA, Phase IV studies are conducted to compare the drug to a competitor, explore additional patient populations, or to further study any adverse events. (From http://www.centerwatch.com/patient/glossary.html [last visited August 23, 2005]).


39. See Dickersin and Rennie, supra note 6.


48. See Abbasi, supra note 43 and Cohen, supra note 45.


51. See Rennie, supra note 50.
Motion(less) *in Limine*

Giles Scofield

“When the two come into conflict, democracy takes priority to philosophy.”
*Richard Rorty*

“There are some people who use philosophy to lead people astray.”
*St. Augustine*

As any seasoned litigator knows, occasionally one interposes an evidentiary objection not simply for the sake of preventing this or that from occurring in court, but also for the purpose of alerting a court to and educating it about the likelihood that it will have to rule on what may prove to be a substantial evidentiary dispute. Instead of waiting until a trial has begun and somebody else’s expert witness is ready, willing, able and – most of all – present to testify, it can be useful to file a motion *in limine*, a pre-trial motion that seeks to bar or set limits to an expert’s testimony.1 Even if the motion *in limine* is not ruled favorably upon at the pre-trial stage, one will have alerted the judge to the nature and basis of the objection, have fleshed out the other side’s arguments, and still be able to renew the objection at trial – e.g., before the witness takes the stand, after the witness’s qualifications have been examined, before the witness offers his or her opinion.

The recently published articles about whether medical ethicists ought to be permitted to testify as expert witnesses suggest that this debate remains *in limine*, and the ethicist’s status in limbo. That the courts either have – in some instances – refused to allow ethicists to testify as expert witnesses and – in other instances – wondered why they bothered to do so, means that the matter is not the “slam-dunk” it was once supposed to be. As a result, the arguments in favor of allowing ethicists to testify display increasing variability, as even those who believe that ethicists ought to testify as experts disagree among themselves about the best way to approach the matter. If we simply thought about the judiciary as if it were a legislature, or about ethicists who are casuists as if they were common law judges, for example, we could and would see the light and allow ethicists to do what they think they ought to do and what they think others need them to do, i.e. testify as expert witnesses and share with others their expert opinion on what ought to be done in a given case.

Instead of shifting around for a better argument, I am simply going to stand my ground and do what one does when one renews an evidentiary objection, which is to say something more about what underlies the objection, delve more deeply into the matter, and comment briefly upon what I take to be the central points others are making. In effect, I am going to proceed as if my motion *in limine* was not granted, which would mean that the medical ethicist has taken or is about to take the stand, and that it is time to renew and re-state the objection.

**Voir Dire**

For starters, I am going to assume what I have always conceded and what is not disputed, which is that there can be and is some such thing as ethical expertise with respect to descriptive ethics and meta-ethics, and even normative ethics to the extent it is something one teaches (as opposed to something one preaches). There is no disputing that certain persons are qualified to teach such topics, and that others are not. Given the claims medical ethicists make on their own

---

Giles Scofield, J.D., is a student at Yale Divinity School.
Given the claims medical ethicists make on their own behalf, it would be fair and reasonable to contend that they are deemed to know what a reasonably prudent ethicist ought to know – notwithstanding that the precise contours of their cognitive domain remains mysteriously problematic absent formal standards for licensing and accreditation.

This is an old question; there are old tensions at work here: between truth and opinion, reason and will, value and preference, the one and the many. These antipodal pairs differ from one another, and none of them quite matches the pair “philosophy and democracy.” But they do hang together; they point to a central problem. Philosophers claim a certain sort of authority for their conclusions; the people claim a different sort for their decisions. What is the relation between the two?

After noting that it has been common for philosophers to be separatist in thought, conformist in practice, Walzer further observed that there are occasions when, instead of being detached from society, philosophers become engaged with it. This occurs when the philosopher finds himself in a position to transform practice into some nearer approximation to the truths of his thought. He cannot be a participant in the rough and tumble politics of the city, but he can be a founder or legislator, a king or nocturnal councilor, or a judge – or, more realistically, he can be an advisor to such figures, whispering in the ear of power.

In other words, the philosopher may opine (or consult – take your pick), an activity that Walzer describes as follows, knowing what he knows about philosophers. Shaped by the very nature of the philosophical project, [the philosopher] has little taste for bargaining and mutual accommodation. Because the truth he knows or claims to know is singular in character, he is likely to think that politics must be the same: a coherent conception, an uncompromising execution.

No matter that philosophers may disclaim any and all interest in politics, there is always the possibility of a partnership between philosophical authority and political power. Reflecting on that possibility, the philosopher may, like Thomas Hobbes “recover some hope
Within the present context, the ethicist’s hope might be that someone on the Supreme Court might read and embrace the position advanced in the amicus curiae brief one has added one’s name to, or refer to and cite the article one has written. Or perhaps some senator or congressional representative – hopefully the powerful chair of a powerful committee – will read the ethicist’s recently published book and invite him or her up to the Hill. Who knows? Maybe the White House will call. In any event, even the chance to participate and testify in ground-breaking litigation would not be such a bad thing. This is what comes of “know something about knowing nobody else knows so well,” which is why bioethicists are plying their trade as...consultants in hospital settings, and as researchers and educators with university affiliations. With increasing frequency, bioethicists are also functioning in a variety of peripheral contexts, such as media interviews, media panels, and public education forums. We find bioethicists on commissions and committees, in the courtroom, in the boardroom, and in the corridors of power – virtually everywhere that expert opinion is sought on the issue of the day.

Being human, all too human, medical ethicists welcome this turn of events. As Laurie Zoloth has written:

We are lovers of knowledge, “philosophers” by nature and name, and we are passionate in particular about secret knowledge, secret languages, codes to which we hold the key. Once one has possession of the inner secret knowledge, one is in demand – sought after as the expert in the field, invited to reflect in prestigious journals, to speak on panels, to be a member of national boards, to be invited into discourse with the media. For this sort of social capital there are no restrictions and no regulations.... It is tricky to turn [down] one’s fifteen minutes of fame, particularly in the American context, where attention is currency. Bioethicists like to be chosen.... It is an ontological issue – our choseness.

Returning to Walzer’s article, one finds he draws one conclusion about these kinds of dreams: “Philosophical founding is an authoritarian business.”

Here’s why.

The claim of the philosopher...is that he knows “the pattern set up in the heavens.” He knows what ought to be done. He cannot just do it himself, however, and so he must look for a political instrument. A pliable prince is, for obvious practical reasons, the best possible instrument. But in principle any instrument will do – an aristocracy, a vanguard, a civil service, even the people will do, so long as its members are committed to philosophical truth and possessed of sovereign power.

Once “the people” enter the picture, however, difficulties arise, the kind of difficulties that have something to do with the “tensions at work here,” the tensions that exist between philosophy and democracy, between the philosopher, who “knows what ought to be done,” and the people.

[The people] may not know the right thing to do, but they claim the right to do what they think is right.

What, in the final analysis, is the problem? That “it is a feature of democratic government that the people have the right to act wrongly – in much the same way that they have the right to act stupidly.” And why is this a problem? Because, to the philosopher, the foundational claim of the democratic way of life makes no sense whatsoever, since “it can never be right to do wrong (not, at least, once we know or can know what is right).”

After an extended discussion, in which he examines the role that judges and the judiciary play in a democracy, as well as the tension that exists between the notion of judicial review and the democratic way of life, Walzer returns to the tension that exists between philosophy and democracy.

Philosophical validation and political authorization are two entirely different things. They belong to two different spheres of human activity. Authorization is the work of citizens governing...
themselves among themselves. Validation is the work of the philosopher reasoning alone in the world he [or she] inhabits alone or fills with the product of his [or her] own philosophical speculations. *Democracy has no claims in the philosophical realm, and philosophers have no special rights in the political community. In the world of opinion, truth is indeed another opinion, and the philosopher only another opinion maker.*

**Recess**

I have quoted extensively from Walzer’s article for several reasons. One is to reveal that, to the extent my objection has something to do with there being some kind of inconsistency between the foundational beliefs of a democratic society and the belief that there is or can be some such thing as “normative ethical expertise,” there is considerably more to my objection than what others routinely reduce it to, which is some vague claim about moral equals and moral equality. What is surprising is not that others try to minimize the gravity of my objection, but that the very persons who profess either to be ethics experts or to know something about ethical expertise evidently do not see the foundational belief of a democratic society for what it is: a normative claim, i.e. a belief about how individuals ordinarily ought to be regarded (as equals) even though they may not be and in all likelihood are not equals. The claim that “all persons were created equal” is not a statement of fact about how the world is, but a statement of belief about how the world ought to be. That’s what makes it a radical claim, for reasons which have been well articulated by Philip Selznick.

The principle of equality is drawn from experience. In embracing it we take into account the evils that ensue when it is rejected or ignored. Although moral equality is a normative idea, it has a factual basis. Most important is that all persons are roughly equal in moral competence and vulnerability. All...persons have much the same capacity for moral choice in their personal lives. We do not expect priests or philosophers to be significantly better than anyone else at making choices. Despite differences in experience, education, character and reflection, no moral elite exists.

This being the case, then if others do not or will not see this for the normative claim that it is, either it is the case that they lack the very expertise they profess to possess; or they agree that the democratic way of life is predicated on the foundational belief I have articulated, but that it can be reconciled with their claim of expertise; or they are simply paying lip-serv-

**Re-Direct (?)**

There is, of course, an obvious objection that can be made to Walzer’s claim, which is that even his opinion is just that; his opinion. The problem with such a response is that it is more of a retort than a refutation, the kind of response that is consistent with and typical of the so-called “analytic style,” the kind of response that is more likely to inhibit dialogue than foster it, to bring the conversation to an end move it along, to score “gotcha” – in this instance *tu quoque* – points and nothing else. There is no denying that some people do believe that there is some such thing as expertise at normative ethics, and that “ethicists” have what it takes to untangle an ethical dilemma.

According to Albert Jonsen:

Ethicists should be able to do this, not because they are insightful and clever, but because they have studied with care the logic and substance of the historical debate over principles and their application. Just as a good statistician can occasionally improve a research protocol by pointing out a flaw in design that a good clinician might miss, so a good [clinical ethicist] should flag the logical flaws in an ethical argument that even good people may not notice.

According to Donnie Self:

There have long been societally recognized moral experts in all cultures....Some people know more about physics than others. Similarly, some people know more about moral values than others. More importantly, some people know more about the process of moral decision-making than others, which is, at least partly, why we have ethics consultants.

According to E. R. Klein:

If...ethical expertise is possible, it would seem that philosophers should be first in line for this honor. Philosophy is a long-standing profession, and philosophers are expert in their field in precisely the same way as other professionals in our society....To really do justice to cases one must have done the philosophical work. Without hubris, it must be concluded that, with very few exceptions, only philosophers have done the requisite work.
Arguably, this kind of re-direct of the proffered expert would be designed to demonstrate, in a manner arguably consistent with Walzer’s essay, that there is another “school of thought” about all of this, one that is consistent with the belief that the ethicist’s voice ought to have an “overriding claim on the attention of others,” for reasons having to do with the ethicist’s supposedly “knowing something about knowing which nobody else knows so well.”

There are several problems with this rejoinder to Walzer’s argument. The first is the same criticism that can be leveled against Walzer, i.e. that it is just their – as opposed to his – opinion. He has his opinion; others have theirs. That being the case, the question remains: What are the rest of us to make of this difference of opinion? That there are simply two schools of thought on the matter? Or that, as between these two opinions, the one reconciles the claim that the philosopher wants to make on us with the claim that democracy makes upon us all, while the other fineses the issue in an Orwellian manner, by paying lip-serv­ice to the notion that each person is the moral equal of every other, except for those who are more morally equal than others?

If Walzer belongs to the democratic school of thought, it is worth asking, notwithstanding Klein’s thoughts to the contrary, whether others can or do believe and say that they do about themselves “without hubris.” In 1984, when applied ethics emerged as a serious endeavor, Annette Baier wondered whether the result might be, not the “plague of gadflies” it was promised to be, but a “nuisance of owls” instead.

[N]o greater hubris is involved in trying synoptically to survey a society’s array of social roles, and its alternatives, than in purporting to produce moral principles valid in all conditions, or truths for all possible worlds....The very professionalization of philosophy makes the likelihood more remote that...[the] awkward questions necessary for a healthy social consciousness [will] come from philosophers.

Even if the ethicist’s claim can be made without hubris, one must ask whether it can be made as ethicists themselves believe it ought to be made, with honesty, forth­rightness, self-knowledge, prudence, and humility.

Another problem is that there is no “school” as such. There are no standards to accredit programs that turn out bioethics experts, no qualifying exams for attaining such a status, no standards for licensure, no code of ethics, no board that disciplines ethicists for un­professional behavior, not even a statute that prohibits the unauthorized practice of this professional field of practice. That there is no way of knowing how an ethicist may be “ordained,” or de-frocked, for that matter, is a state of affairs that exists not as a matter of happenstance, but as a result of the intentional, vol­untary, knowing, and deliberate choice that professional ethicists have made.

In 1993, John Fletcher, who believed that “the time for a laissez faire” had past, wrote that he was especially concerned “about the primitive state of standards and due practice for ethics consultation, as well as the lack of an adequate approach to [the] training and educating...of those who will be consultants. I am confident...that there are many ‘ethics disasters’ waiting to happen. Do we have to wait for these worst case scenarios [before] taking some action?” The field of bioethics did respond to this situation. According to its self-description:

The rapid growth of consultation and concerns about accountability and quality assurance led two major bioethics organizations, the Society for Health and Human Values and the Society for Bioethics Consultation, to convene a national task force to explore core competencies and related issues in ethics consultation....The sponsoring organization of the Task Force selected the co directors, who then selected 19 scholars representing diverse fields and a wide variety of organizations and institutions. A comprehensive bibliography of the empirical and conceptual literature on ethics consultation was compiled. Information was also solicited from bioethics education programs through a national mailing. The Task Force held six meetings in the course of 2 years. Subcommittees developed drafts on specific content areas. Feedback on an initial draft report was obtained from the bioethics community. Revisions were made and a final draft was unanimously approved by the Task Force. The Task Force Report was later adopted by the American Society for Bioethics and the Humanities.

The immediate outcome of these efforts? That the time for a laissez faire approach had not passed, and
that it would be undesirable, imprudent, and counterproductive either to require the certification of individual practitioners or the accreditation of the programs that train and educate them.\textsuperscript{32} And the long-term outcome? That the field finds itself at a crossroads (if not in the crosshairs).

Clinical ethics, like the broader field of bioethics from which it emerged, is at a critical crossroads in its development, with conflicting paths ahead. It can either claim its distinctive place in the clinical arena, insisting unapologetically on certain minimal standards of training, practice, and competence, addressing head-on debates about various models and methodological approaches to consultation, and establishing a shared vision of the purpose and meaning of the enterprise of clinical ethics consultation. Or it can evolve into a hobby that untrained, albeit interested and generally well-intentioned individuals can dabble in for fun and even profit, as they see fit, and without regard to the deep history and rich disciplinary roots in the field, the serious debates in the academic literature of bioethics, the foundational case histories and legal theories, or even any sense of professional accountability. At a time when 95% of those doing ethics consultation in U.S. hospitals have no formal education or training, when there is little or no consensus about how clinical ethics ought to be practiced and by whom, the choice for the future of clinical ethics is a stark one, and never has it been more pressing.\textsuperscript{33}

Understandably, some ethics experts would probably be as reluctant to embrace the opinions of Jonsen, Self, or Klein as they are to embrace that of Walzer. But if there is a third way of talking or thinking about ethical expertise, it has yet to reveal itself in a cogent, comprehensible definition of what it is that this kind of “expertise” is really all about, which may have something to do with why ethics experts are having a tough time having courts accept the assertion that they are experts, and why the field has not yet approached a legislative body with the request that it enact some kind of law that would recognize this “profession.”

Finally, there is the practical problem created by those whose opinion differs from Walzer’s, which is how anyone other than an ethicist can evaluate or pass judgment on the claims that the ethics expert makes. In Walzer’s world, we have the right to embrace or reject the ethicist’s claim, notwithstanding that one such choice might be an instance in which we exercise the right to be wrong. In the ethicist’s world, it also cannot be the case that there is some such thing as ethics expertise. Why? Because, unless it is the case that we are to suspend the prohibition against self-contradiction where ethicists are concerned, then it is absurd to believe both that others do not know enough about how one ought to make evaluative judgments, and that those same others somehow possess enough moral and ethical capacity and insight to approve of there being some such thing as ethical expertise. They cannot know and yet not know at the same time, lack and possess the capacity for making such a judgment at the same time. How can the ordinary person know what he or she ought to make of the claims being made by a “Wizard of Oughts?” If such a person can and may do so, then does not the Wizard cease being the Wizard? By the same token, if the ordinary person can only believe that the Wizard of Ought is some kind of expert, but not know nor comprehend what that expertise consists of, what weight can or should an ethicist honestly place on a public affirmation of something about which the public does not and cannot know?

If one of the problems raised by “democratic” objection to the ethicist’s assertion of normative expertise is that ethicists do not seem to recognize a normative claim for what it is – the notion that all persons ought to be regarded as equal is a normative claim – another problem in this context is that ethics experts may not be as skilled at analyzing and resolving ethical dilemmas as they profess to be. Although they acknowledge that a tension of some sort exists between the notion that the ethicist is an expert at normative ethics and the democratic belief that all persons ought, as a general rule, be regarded as moral experts, they have yet to find a way to resolve this dilemma, but instead seem intent on both concealing the nature of this conflict and on trying to finesse it. Given that so many of the conflicts they involve themselves with consist of tragic choices, it does not bode well that they rely as heavily as they do on subterfuges – such as the ethics facilitation model – and other intellectually dishonest methods of analysis and conflict resolution.

Although it is impossible to predict what kind of impression will be created in the mind of a judge (or the minds of a jury) once information such as this is elicited from a proffered “ethics expert,” it is difficult to believe that it could or would create a favorable impression, of a field that has its act together, especially given that there are no studies that validate the work that ethicists do when they consult,\textsuperscript{34} which is, of course, precisely what an ethicist is doing when he or she testifies as an expert.
A Jury of One’s Peers

Helpful though it is to bring Walzer and Selznick into the debate,35 I am familiar enough with the peregrinations of this debate to know better than to rest my objection on insights drawn primarily from outside the “field” of bioethics. Whether others would prefer additional or different support for my objection, the bioethics literature itself, especially with respect to the importance ethicists have placed on informed consent, provides it. According to Jonathan Moreno,

[T]he informed consent doctrine promoted by biomedical ethics is the lingua franca of modernity. In classical liberal terms, respect for individual self-determination is the very condition for civil society, for entrance into the social contract. The bioethical consensus that has grown up around the doctrine of informed consent in health care is really a reflexive confirmation of our society’s essential philosophy: the answer to the “who should decide” problem is that everyone should decide for himself or herself under conditions of freedom and equality.36

These sentiments echo those previously stated by the President’s Commission:

More is involved in respect for self-determination than just the belief that each person knows what’s best for him- or herself. Even if it could be shown that an expert (or a computer) could do the job better, the worth of the individual, as acknowledged in Western ethical traditions and especially in Anglo-American law, provides an independent – and more important – ground for recognizing self-determination as a basic principle in human relations, particularly when matters as important as those raised by health care are at stake.37

The informed consent doctrine is predicated on the assumption “that many patients are capable of comprehending what they need to know in order to decide what is best for them and that, therefore, they must be treated as adults possessed of the capacity for self-determination,”38 which reflects and corresponds to the notion that they ought to be regarded as moral equals. Additionally, the informed consent doctrine protects and respects the patient’s right to make what others might regard as a wrong-headed, perhaps even a “stupid” decision.39 Finally, the logical corollary to the belief that we protect and respect the moral autonomy of individuals is that the beliefs and opinions of professionals ordinarily should do not outweigh or usurp those of the individual.40

If this is not enough to persuade others that there is something, and not nothing, to this objection, then one need only read beyond and look behind the SHHV-SBC Task Force Report. At the Task Force’s very first meeting one of its members, Sociology Professor Charles Bosk gave a presentation on the sociology of professions, in which he discussed the ways in which professions formerly and currently are analyzed. As part of that discussion, Bosk referred to the “tension” that can exist between claims about expert knowledge and democratic principles.

Where all this comes in for bioethics consultation is in the need to recognize that there is something inescapably elitist, anti-democratic, and perhaps self-serving about writing standards, minimal qualifications for practice, and the like....[T]he tension between expert knowledge and democratic principles has been a recurrent theme since we sociologists began writing about professions. Indeed, one of the most provocative statements of this theme is found in Weber’s “Science as a Vocation,” first published in 1918. In a discipline such as bioethics, in which autonomy is a core value, these questions are particularly acute.41

At its next meeting, the Task Force heard from its Executive Director, Mark Aulisio, who had this to say about the “tension” Bosk had referred to.

In portraying the various models [of ethics consultation], Mark underscored a fundamental dilemma that will have to be addressed if ethics consultation is to be justifiable in our society. Ultimately, ethics consultation must be characterized in a way that is consistent with the basic values of a liberal democracy. The liberal democratic framework within which ethics consultation must be practiced in our society will form part of the framework within which an acceptable normative characterization of ethics consultation must be situated.42

At a subsequent meeting, the minutes indicate that a question that remained open was whether “ethics consultation [can] be understood and practiced in a way that is consistent with the fundamental values of a liberal democracy?”43 Although there is nothing in the minutes that explicitly reflects how the Task Force decided to resolve the matter, there is no denying that its members thought it a serious question, serious enough for some its member later to observe that if ethicists could not consult in a manner that is consistent with the “deep societal values of our liberal democracy...then not only should no set of education-
al and training standards...be adopted, but the practice...should be abandoned.44

With this kind of foundation for the objection, and with the burden of proof resting on those who wish both to exempt the ethicist from the confines of Rule 701, which bars opinion testimony as a general rule, and situate the ethicist within the exception to the rule that is Rule 702, which allows for opinion testimony under closely regulated conditions, there is a reasonable basis both for interposing the objection to the ethicist’s being allowed to testify as an expert, and for believing that this objection will be sustained, not overruled.45

Court Resumes
Because one should expect the best, but anticipate the worst, one must be prepared for the possibility that the ethicist may be allowed to testify as an expert. It is one thing for the court to determine whether the ethicist may testify, and something else for it to determine what the ethicist may testify to and how the ethicist may testify. Presumably, the reason a litigant retains an expert is for the value attached to his or her opinion on the “ultimate question,” i.e. his or her opinion on what ought to have been done. Before one gets to the question of whether the bioethicist may offer an opinion on the “ultimate question” before the court, one must address questions about how this expert may go about the business of communicating with the court in a manner that is educational and informative.

Here is where one is likely to find common ground between concerns courts have about the proper role of experts and concerns ethicists have about how they may and may not “consult.” On the one hand, the ethics literature is replete with statements to the effect that ethicists do not want to impose their values on others or to “apodictically” determine what ought to be done in a given case. According to John Fletcher,

The views of experts in medical ethics must not be allowed to dominate the process of decisionmaking or the outcomes, lest the open commonwealth of values...[become] a closed province of a “tyranny of principles” that shuts down the debate or imposes rigid solutions before examining the facts of the case.46

According to Pellegrino, Siegler, and Singer,

The ethicist’s aim should be to enable, empower, and enhance the decision-making capabilities of [others]...The ethicist should not make the decision for [others], or, worse still, impose his or her ethical values or beliefs on others.47

By the same token, courts are not interested in having expert witnesses usurp the authority of the judge or the jury, hence the common belief that ethicists ought not overreach.

Setting limits to how the ethics expert may testify is entirely consistent with the views held by those who believe, for example, that medical ethics consultants should not adopt an “authoritarian” stance when working with others, but that they should instead employ something referred to as the “ethics facilitation approach.”

The ethics facilitation approach recognizes the societal boundaries that exist for morally acceptable solutions. In contrast to the authoritarian approach...[i]t respects individuals’ rights to live by their values, by not displacing moral decision-making authority or representing the personal views of the consultant only...The ethics facilitation approach is consistent both with the pluralistic context within which ethics consultation is done and the rights of individuals to live by their own values, recognizing that there are definite boundaries within which decisions must be made and helping to ensure that these boundaries are not transgressed.49

828 THE JOURNAL OF LAW, MEDICINE & ETHICS
Because ought implies can, one needs to ask what basis there is for believing that ethics experts can and will do what they ought to do, i.e. communicate basic information about descriptive ethics and meta-ethics, in a manner that does not, cannot, and will not create a risk of their unduly influencing others. As it turns out, there is little reason to believe that they can or will be what they profess to be, i.e. a non-influential influence on how or what others think about a given case. In order to understand both why this is so and why medical ethicists themselves know this to be so, we need to take a brief look at an analogous field of practice with which medical ethicists are familiar: genetic counseling.

Genetic counseling refers to the process whereby families and individuals receive professional assistance that is designed and intended to help them learn about and understand human genetics, as well as clarify the nature and consequences of the treatment options available to them, in order for them to integrate this information into their value systems and choose as seems best to them. Given both the history of genetic interventions and the often controversial and highly personal nature of the choices individuals and families must make, the desirability and necessity of preserving and protecting individual autonomy is especially acute. This has led genetic counselors to see if they can communicate important information to others in a manner that does not encroach upon, usurp, or otherwise unduly influence their ability to make a knowing, autonomous and authentic choice. The solution they have arrived at is referred to as “non-directive counseling.”

According to the Encyclopedia for Bioethics, counselors want to avoid, to the extent possible, being accuse of “playing God,” and to resist any temptation to practice eugenics the process of manipulating genes in order to “improve” the genetic makeup. [Such] manipulation is accomplished by directing the counselees about what reproductive decisions they should or should not make. This is in appropriate because respect for autonomy should be a predominant ethical values guiding the counseling process and its outcome.

Non-directive counseling is and remains the “gold standard” among genetic counselors. According to Aubrey Milunsky,

It is a fundamental principle that genetic counseling be non-directive and that physicians desist from visiting upon patients their [own] religious, moral, ethical...or other prejudices.

According to John Fletcher and Dorothy Wertz,

Nondirective counseling means that counselors refrain from giving moral advice, even if asked for it, out of respect for the moral autonomy of patients.

According to Leslie Biesecker, “the nondirective standard protects [others] from the biases of medical professionals and [from] other pressures.”

As the Institute of Medicine explicitly emphasizes,

The principle of autonomous decision making requires that providers not present any reproductive decisions as “correct” or advantageous for a person or society.

Laudable though the gold standard of non-directiveness is, ours is not a perfect world. The line that distinguishes directive from non-directive counseling is not well defined, and may be impossible to define with precise certainty. Moreover, it has proven to be one thing to espouse the ideal of non-directiveness, and something else to adhere to it. Finally, even those who try to be non-directive may not always succeed at doing so, because of what others may read into or out of what the counselor says or does not say, or into and out of how the counselor says what he or she says. As a result, one expert has said “ostensibly non-directive counseling...is a sham, not because of a personal failure on the part of the counselor, but as a direct result of the encounter between the counselor and the client.”

In short, it is virtually impossible to engage in non-directive counseling without running some risk of subverting or unduly influencing the moral autonomy of others. The risk comes with the territory, as every counselor (and every teacher or educator) knows.

Abbreviated though this discussion of genetic counseling is, it enables us to return to the questions at hand, and better assess whether, to what extent, and if so, how, an ethicist ordinarily may testify as an expert witness, assuming he or she is in fact allowed to do so. For one thing, it is impossible for ethicists to be “value-neutral.”

It is impossible for ethics consultants to be value-neutral. Consultants typically will have their own moral views about the issues [present] in case consultations and about how cases should be best resolved. These views will inevitably influence their...work.

For another, it is difficult to know when ethicists are merely facilitating the decisionmaking process others
are engaged in and when they may be doing something more than that.

The line between guiding and driving the discussion is very difficult to draw, much like the line between persuasion and manipulation in informed consent discussions.62

For reasons such as these, ethicists “need to be sensitive to this and not usurp moral decision-making authority or impose their values on others.”63

Given these kinds of considerations, not only would it seem to be the case that one may properly object to an ethicist’s rendering his or her opinion on the “ultimate question,” it would also seem to be the case that an honest ethicist could not and would not object to such an objection. In fact, were a judge to ask for or insist upon knowing the ethicist’s opinion, the ethicist ought to say that it would be improper for him or her to say. The same considerations that preclude the ethicist from saying something about what ought to be or ought to have been done in a given case dictate that, should the ethicist be permitted to testify on matters pertaining to descriptive ethics and metaethics, that his or her testimony be closely regulated, so that he or she does not cross the line, and start preaching under the pretext of merely teaching.

What Never?

“Well, hardly ever…”

Gilbert and Sullivan, H.M.S. Pinafore

I do not mean to suggest or imply – or for others to infer – that there are no conditions under which a medical ethicist might be permitted to testify as an expert witness, even to the point of being permitted to offer his or her opinion as to what ought to be or (more likely) to have been done. I just don’t think it’s going to happen in accordance with any of the various scenarios hypothesized by those who think that it will and believe that it should, i.e. a case in which everyone is wringing their hands, because they do not know what to do about this or that ethical dilemma that confronts them, until a medical ethicist appears on the scene, cuts the Gordian Knot, and rides off into the sunset, while everyone wonders just who that masked man (or woman) was.

Instead, I think the most likely scenario will be one in which a medical ethicist is being sued for malpractice, i.e. for having failed to do what a reasonably prudent medical ethicist would and should have done.64 While I fully expect that something analogous to a “conspiracy of silence” will make it difficult for an attorney to find an ethicist who will testify that another ethicist breached the professional standard of care, some day one brave soul will dare provide such testimony. And what will happen then? In all likelihood, many of the arguments I have raised against allowing ethicists to provide expert testimony – arguments which, as is clear from the recently published and other articles, others find so disagreeable – will be used to prevent a proffered expert in medical ethics from testifying against another medical ethicist from testifying as an expert, or to negate the impact of his or her testimony, thereby demonstrating that, in the final analysis, it all boils down to whose ox is being gored, to whether the field of medical ethics will benefit from or be harmed by whether, and if so how, a medical ethicist testifies as an expert witness.

And at that juncture, a court overseeing this interesting case of first impression, in which a medical ethicist is being sued for a failure to conduct him or herself in a professionally responsible manner, will, not quite knowing what to do, derive some guidance from the articles penned by Imwinkelried, Spielman, Latham, Nelson, Majunder, and Kipnis. Confronted with such a situation, a court could in the course of exercising its judicial function, assume its “legislative” role, and do what the field of ethics ought to have done by now, but which it has failed to do: articulate what standard of practice, otherwise known as the field’s standard of care, ought to be. Because the court would be acting in its legislative capacity, the “technical evidentiary standards [would be] inapplicable.”65

Under such relaxed conditions, the court might allow an ethics expert who might otherwise have been precluded from testifying do so, thereby overcoming both the vicious objections being raised on behalf of the defendant-ethicist and circumventing the “conspiracy of silence” that is likely to come into existence if and when an ethicist is finally sued. The judge could then conclude, perhaps even take judicial notice of a problem identified by Spielman and Nelson, which is that “the field of bioethics [needs] to develop and actively implement its own guidelines for reducing the current chaos in forensic bioethics work.”66 The court could also conclude, again from what Nelson has to say, that the field of bioethics lacks “professionalism,” because it has failed to get its house in order. On the one hand, “no legal license or social mandate exists that can be used to identify a bioethicist, as there is for a physician, a general contractor, or even a cosmetologist.”67 On the other hand, “bioethics does not meet all of the standard sociological criteria for a ‘profession.’”68

From Latham, the court could conclude that to the extent that ethicists act like casuists, they are acting like judges69 – except, of course, for the facts pointed out by Spielman, which are that they lack many of the
features that enhance the law’s reliability, of which one would certainly be the standardized and public presentation of the work that bioethicists do, the “cases” that they participate in. From Kipnis, the court could conclude that it would be reasonable for it to rely on consensus documents and a survey of the literature in order to determine what the standard of care ought to be. Finally, from Majunder, the court could conclude that ethicists believe that accountability matters, and that it should be the rule and not, as in the case of managed care, the exception to the rule.

With these articles in hand, and with the benefit of such other fact-finding as the court chooses to undertake, it will be able (and presumably willing) to do what the field of bioethics is able, but evidently unwilling to do, which is to transform this informal, of profession into a formal, sub jure, if not a de jure one, thereby restoring the rule of law and reaffirming the priority of democracy to philosophy. It is difficult to see how anyone – including Majunder, Kipnis, Nelson, Spielman, Latham, and Imwinkelried – could or would honestly object to such a result.

Ironically, by demonstrating the extent to which the field of bioethics operates informally, these authors have revealed the extent to which ethicists operate outside and beyond the reach of the law, as a wholly unregulated field, not even a self-regulated one. Tragically, it is the field’s failure to take professionalization seriously, to think critically and self-critically about the practice, that will, in the final analysis, cause it to experience the “ethics disaster” of having regulation imposed upon it by others, if only to prevent the unregulated in-court statements of “ethics experts” from doing as much harm as their indiscriminate out-of-court utterances already have done.

References
1. See, for example, in Re Rezulin Products Liability Litigation, 309 F. Supp. 2d 531 (S.D.N.Y. 2004).
4. Id. at 379.
5. Id.
6. Id.
7. Id. at 381.
8. Id.
9. Id. (quoting Thomas Hobbes).
14. Id. (emphasis added).
15. Id. (emphasis added).
16. Id. at 385.
17. Id. at 387.
18. Id. at 397.
19. Nor is it the case that Walzer is the only “authority” I can cite in support of my objection. “[F]ew philosophers, and probably not many thoughtful and educated people, now believe that we can arrive at absolute, intersubjectively valid, and ‘objectively true’ moral judgments…Although some more philosophers may make such a claim, they have conspicuously failed to demonstrate the absolute and objective status of any specific moral judgments they are prepared to assert. Instead, their ‘objective moral truths’ turn out to be highly debatable [and ] their pre-tense of intersubjective validity cannot be upheld.” R. A. Dahl, Democracy and Its Critics (New Haven, CT: Yale University Press, 1989): 66; see also, E. Sagan, The Honey and the Hemlock: Democracy and Paranoia in Ancient Athens and Modern America (Princeton, NJ: Princeton University Press, 1991).
21. The last of these scenarios being consistent with a concern that Eliot Freidson voiced about the professions years ago, when he wrote, “there is a real danger of a new tyranny which sincerely expresses itself in its own language of humanitarianism and which imposes its own values on others for what it sees to be their own good.” E. Freidson, Profession of Medicine: A Study of the Sociology of Applied Knowledge (New York, NY: Harper & Row, 1970): 381.
32. Id. at 66-67.
39. ["The individual person, as the one most vitally concerned, is allowed to choose... even though he [or she] may elect the foolish and disastrous course." H. W. Smith, "Antecedent Grounds of Liability in the Practice of Surgery," Rocky Mountain Law Review 14 (1942): 237."
42. SHHV-SBC Task Force on Standards for Bioethics Consultation, Meeting Two Minutes: December 13-14, 1996, Final Version (undated) (prepared by M. Aulisio, Ph.D., Executive Director).
43. SHHV-SBC Task Force on Standards for Bioethics Consultation, Common Themes and Disputed Questions from the Literature Review and Essay Exercise, December 13, 1996 (Prepared by M. Aulisio, Ph.D., Executive Director).
45. See, for example, In re Rezulin Products Liability Litigation, 309 F.Supp. 2d. 531 (S.D.N.Y. 2004).
49. Id. at 61.
50. Id.
56. Institute of Medicine, supra note 51, at 171 (emphasis in origin- nal).
57. "One might think that directiveness and non-directiveness are polar opposites on a single dimension. I would like to suggest that they are not. In fact, there are more similarities in the two approaches than meets the eye." S. Kessler, "Psychological Aspects of Genetic Counseling: Thoughts on Directiveness," Journal of Genetic Counseling 1 (1992): 9-17, at 9.
58. "Professional commitment to a value-neutral, non-directive style guides practice, [but] should not be confused with prac- tice....Value-neutral, non-directive counseling is...easy to espouse, but in practice difficult to perform." C. L. Bosk, All God's Mistakes: Genetic Counseling in a Pediatric Hospital (Chicago, IL: University of Chicago Press, 1992): 153.
60. "It is almost impossible to avoid communicating, by tone of voice, subtle changes of expression, choice of words, and even what one does not say, some signals that are translated by [others] as directive." A. Clarke, "Introduction," in A. Clarke, ed., Genetic Counselling: Practice and Principles (London: Routledge, 1994): 11. "There are...implicitly directive compo- nents to the counseling relationship conveyed through body language, emphasis, and time spent." Murray, supra note 52, at 925.
62. Id.
63. Id.
D. N. Sontag, “Are Clinical Ethics Consultants in Danger? An Analysis of the Potential Legal Liability of Clinical Ethicists,” University of Pennsylvania Law Review 151 (2002): 667-705. By citing Sontag, I do not mean to endorse his analysis. Whereas Sontag’s argument is that ethicists ought to bear little responsibility because their involvement tends to be de minimis, I would argue that because ethicists wish to minimize their responsibility, i.e. maximize their irresponsibility, that their involvement should be circumscribed and minimized to the greatest extent possible. Quite frankly, I am surprised that no one seems to think that the professionalization of the field of ethics, of which establishing the ethicist’s status as a legal expert is one component, raises no anti-trust concerns. But that is another story, for another time.

Imwinkelried, supra note 2, at 211.


Id. at 263, n.60. Nelson’s assertion that not “just anyone can claim to be an expert at bioethics…or to do bioethics in a professional manner and not as an amateur or hobbyist,” id. at n.61, must be contrasted both with the concerns voiced by Rubin and Zoloth, see note 33 and accompanying text, supra, and with the matter of fact statement made in connection with a recently published study of whether ethicists can help reduce health care costs in the ICU. T. Gilmer, L. J. Schneiderman, and H. Teetzel, et al., “The Costs of Nonbeneficial Treatment in the Intensive Care Unit,” Health Affairs 24 (2005): 961-71. According to the authors of that study: “Ethics consultations were provided...by people equipped with medical, doctoral or law degrees; by social workers and theologians; by those formally schooled in ethics and philosophy; and by those who had acquired their expertise one way or another during the course of their career.” Id. at n. 8 (emphasis added). Interestingly, the informed consent of these patient-subjects was not obtained, it being deemed not to be part of the standard of care. Id. at n.9; but see, R. M. Veatch, “Terri Schiavo, Son Hudson, and ‘Nonbeneficial Treatments,’” Health Affairs 24 (2005): 976-79; A. H. M. Antonmmaria, “Do as I Say, Not as I Do: Why Bioethicists Should Seek Informed Consent for Some Case Studies,” Hastings Center Report 34, no. 3 (2004): 28-34.

Id. at 263, n.61.


According to Jonathan Moreno, “were critical theorists to write about the institutionalization of bioethics...they could truly have a field day.” J. D. Moreno, “Can Ethics Consultation be Saved? Ethics Consultation and Moral Consensus in a Democratic Society,” in Aulisio, Arnold and Youngner, supra note 35, at 32.

These are indeed dangerous times. In the name of “cost-effectiveness,” we cut back health benefits to the poor, who are more likely to be sick than the nonpoor. We miss our chance to heal. In the setting, we’re told, of “scarce resources,” we imperil the health care safety net. In the name of expedience, we miss our chance to be humane and compassionate.1

Medicaid is again – still – the subject of reform discussions in Washington and in state capitals. The program has been subject to varying, sometimes conflicting pressures since its inception. Its primary purpose has been serving the health care needs of the poor and disabled. It was structured, however, to appeal to (or at least to not alienate) private health care providers. In addition, its mix of state and federal funding and control was intended to draw in the states as partners.2 As it now exists, Medicaid finances health services for over 50 million of the most vulnerable Americans, including poor children and their families, the elderly in nursing homes, and people with disabilities; in America’s strained health finance universe, Medicaid fills many gaps left by a largely market-based system. It covers one-third of all births, it funds over 40 percent of nursing home days, and is the dominant insurer for care for people with serious mental illness and HIV disease.4 Current reform efforts are focused on cost reduction. The proposals range from minor adjustments in vendor payment methodologies to sweeping structural shifts in federal-state control over the program. Short term budget savings are likely to be in the forefront this year, although significant structural changes are clearly on the minds of many governors and some members of Congress. Evaluation of these potential reforms should focus on two questions: will they signal a lessened commitment to Medicaid coverage for the poor and vulnerable, and, if so, in favor of what alternative?

Why Reform Proposals Now?
The beneficiaries of Medicaid could make substantial arguments for reform of the current program. In recent years, for example, groups of recipients have sued state Medicaid agencies alleging the failure to provide appropriate mental health services; the elimination of coverage of necessary therapeutic services for special needs children; and failure to make available necessary children’s dental services.5 In many states, substantial procedural barriers to enrollment have arisen, apparently as a means to reduce the number of Medicaid recipients.6 In some states – Missouri and Tennessee, for example – drastic cutbacks in Medicaid programs have left thousands without coverage.7 With the federal courts increasingly hostile to lawsuits from Medicaid beneficiaries,8 beneficiaries and their advocates could have been the movers behind the reform efforts. They are not.

The current reform movement is driven by state governments. Medicaid has long been a major part of state budgets. In the 1990s, when states benefited from very strong economic conditions, they were able to cut taxes and fund Medicaid budgets as well. As the economy turned sour, state officials experienced increases in Medicaid beyond the rate of background inflation, and well beyond the rate of their revenue increases.9 The increases in Medicaid costs and declines in state revenues arose in a period of anti-tax

John V. Jacobi, J.D., is Professor and Associate Director, Health Law & Policy Program, Seton Hall Law School

Dangerous Times for Medicaid

John V. Jacobi

834 THE JOURNAL OF LAW, MEDICINE & ETHICS
movements, inclining state officials to equate Medicaid reform with cost containment. This cost containment manifested in restricted enrollment through administrative changes, reductions in the scope of benefits, increased beneficiary cost sharing, and reduced reimbursement rates. More extensive cost-saving changes required either waivers of federal requirements or statutory changes at the federal level, however, and the states began to advocate for Medicaid reform.

The Bush administration was receptive to calls for changes in Medicaid. The President called for cuts in the program, and in April 2005, Congress adopted a budget resolution that included a plan for $10 billion in Medicaid spending in the next five years. HHS Secretary Leavitt created a Medicaid Commission charged with proposing the reforms that would achieve the five-year goal of reducing Medicaid spending by $10 billion, and proposing longer-term reforms to “ensure the long-term sustainability of the program.” Following the release of the Commission’s report on five-year reforms, Congress has begun to work on drafting budget language consistent with its earlier budget resolution achieve $10 billion in cuts, work that has been complicated by responses to hurricane Katrina and several political factors. In the meantime, The Centers for Medicare and Medicaid Services continue to grant states waivers from federal Medicaid requirements; Florida was recently granted a sweeping waiver that will enable it to reshape – and reduce – the benefits available to Medicaid beneficiaries. The general shape of Medicaid reform from the perspective of the states, the administration, and Congress is clear: reform means reduced spending.

The Range of Proposed Medicaid Reforms

Reduced cost is the goal of the reform efforts in Congress. The exact shape of the spending reductions is unclear at this point for several reasons. First, reforms are proceeding along two parallel paths: “short-term,” with the goal of reducing the federal Medicaid expenditures by $10 billion over the next five years, and “long-term,” with the goal of “ensuring the long-term sustainability of the program.”

Another axis of complexity is the slightly different orientations of the state and federal governments in this reform effort. For states, the goal is increased state control over benefits and eligibility design without loss of federal matching funds, while Congress appears to be focused on the bottom line goal of reducing federal spending. The reform efforts will likely realize some cost-reducing changes in 2005, but the time horizon is much broader, and promise continued pressure for change in coming years. The proposed reforms can be fit into four categories.

1. Paying less for drugs. As the National Governors Association says, “[s]tates and the federal government have long suspected that Medicaid overpays for prescription drugs.” It is suspected rather than known because the means by which CMS receives price information from the pharmaceutical companies creates assessment and audit difficulties. One reform in this area, whether or not there are reductions in prices paid for drugs, is likely to be improved guidance from CMS on pharmaceutical firms’ reporting, the collection of data by means that more readily permit auditing, and increased regulatory oversight of drug rebate and pricing programs. The cost of drugs itself, however, is likely to be addressed directly. It is likely, for example, that the pricing of drugs for Medicaid payment will shift from Average Wholesale Price (“AWP”), a notoriously murky standard and one that is believed to be routinely inflated, to some other standard, such as “average manufacturer’s price” or “average sales price” to be subject to more readily verifiable standards. In addition, the Medicaid Commission has recommended permitting the Medicaid rebate to be available to states when outpatient drugs are accessed through a Medicaid managed care firm, a gap in the current rebate provisions.

2. Seniors, long-term care, and “poverty” programs. The cost-cutting move that fires Medicaid reform in Congress is, as is described above, as much about cost-cutting as it is about reform. For that reason, the emphasis in Congress and elsewhere is sometimes more on reducing the federal budget than it is on reviewing Medicaid programatically; in fact, this budget-cutting focus has lead some in Congress to speak of the need to cut $10 billion from Medicaid, and some of the need to cut that amount from Medicaid and Medicare combined. There is a strong feeling among some members, particularly in the House of Representatives, to focus on Medicaid rather than Medicare, presumably because middle class

As it now exists, Medicaid finances health services for over 50 million of the most vulnerable Americans, including poor children and their families, the elderly in nursing homes, and people with disabilities.
retirees (the beneficiaries of Medicare) are a greater political force than are the poor and low-income workers (the beneficiaries of Medicaid). One reform proposal that is likely to raise the ire of middle class retirees and their advocates is that to make it more difficult to seniors (and others) to protect assets while becoming eligible for nursing home care. One would increase the “look back” period prior to an application for Medicaid during which states could examine asset transfers for penalty purposes from three to five years. Another would permit the penalty for a transfer within the look-back period to begin to run from the date of Medicaid application rather than the date of the transfer – a change that could significantly affect seniors entering nursing homes. A third reform would require that seniors use the equity in their homes (through reverse mortgages) to finance nursing home care as a condition of gaining eligibility for Medicaid. Taken together, these changes could sharply limit the ability of many middle class seniors to shift assets to easily obtain Medicaid funding for nursing home stays.

3. Personal responsibility. A major aspect of many Medicaid reform proposals emphasizes the need to create incentives for recipients to be more prudent purchasers of care. Many conservatives seeking health care cost containment focus on the need to address the “moral hazard” that allegedly results from the insulation of health care consumers from the cost of their care by imposing cost sharing in the form of co-payments and co-insurance. The purpose of imposing this cost sharing is not to gain the income from the relatively small amounts that would be paid by beneficiaries. Instead, it is to create incentives for beneficiaries to refuse care they would otherwise receive in the absence of cost sharing, or to accept less costly care. As the National Governors Association put it,

States should be given the ability to implement common-sense, enforceable cost-sharing throughout the Medicaid program both to increase responsibility of Medicaid beneficiaries for the cost of their health care, and encourage cost-effective care in the most appropriate setting.

The same point was put more tersely by a Republican member of the House: “If people have a personal stake in the cost of their care, they will use it more responsibly.”

The cost sharing proposals include increasing the use of co-payments at the time of service, particularly for those beneficiaries with family incomes above 100% of the federal poverty level (that is, more than $16,090 for a family of three). In addition, some reform proposals emphasize that cost-sharing should be “enforceable,” which means that health care providers should be permitted to refuse services for those unable to pay the copayment. (“Enforceable” cost sharing schemes permit the refusal of medically necessary care to beneficiaries unable or unwilling to pay their copayments or deductibles.) Cost-sharing proposal would create “tiered” copayments for prescription drugs, which would create “enforceable” charges varying depending on whether the drug prescribed for the patient is a “preferred” or a more costly drug. A third proposal would broaden states’ ability to charge premiums for Medicaid, denying otherwise eligible Medicaid recipients any coverage unless they paid an annual premium.

4. Restructuring and flexibility. A long-term goal of many would-be Medicaid reformers is the restructuring of the benefits and eligibility aspects of Medicaid, and the extension to states of more power to restructure benefits and eligibility. One of the reforms would permit some Medicaid recipients to be eligible for the reduced benefits permitted under the State Child Health Insurance Program (“SCHIP”) rather than the richer package of benefits required under existing Medicaid law. SCHIP benefits can be based on commercial insurance products, and omit coverage of many of the services now available to Medicaid beneficiaries. Advocates of restructuring also advocate a reorientation of Medicaid from a fixed menu of covered services to a program geared to achieve positive health outcomes through, for example, the use of disease management programs and centers of excellence.

A second-order restructuring recommendation would permit states flexibility to make benefits and eligibility changes at their option, or pursuant to an easier to satisfy waiver process. Medicaid program structure would then depend less on Congressional judgment reflected in federal statutes, and more on shifting discretionary judgments by governors and state legislators. The recent approval of a Florida’s Medicaid waiver application may give some indication of the uses to which such flexibility may be put. The Florida waiver appears to be a variant of a “defined contribution plan” for many Medicaid recipients. Medicaid eligibility currently entails entitlement to a defined range of benefits of “sufficient in amount, duration and scope to reasonably achieve its purpose.” Under the Florida waiver, some beneficiaries will no longer be entitled to federally-defined benefits, but will instead receive a contribution of a set value to purchase coverage from participating plans. These plans need not cover all Medicaid-mandated
services, although they must be “actuarially equivalent” in value to Medicaid coverage. It is unclear whether any residual Medicaid coverage will be available to those who need services uncovered by these plans. In addition, beneficiaries will be permitted to “opt out” of the Medicaid system altogether, receiving a voucher that can be applied to coverage from another source, for example, their employer.36

A set of reform proposals seeking to reduce federal costs comprise a federal counterpoint to the states’ desire for flexibility. The federal government would like to limit states’ ability to draw down creatively on federal matching shares for new or expanded programs. Tension has long existed between the states and the federal government over states’ use of inter-

governmental transfers and other “accounting tricks” to increase federal Medicaid expenditures without any genuine match from the states. The administration’s legislative proposal for Medicaid reform included restrictions on such moves.37 In addition, the Administration’s reform proposals include a recommendation to convert disproportionate share hospital (“DSH”) payment – the payments that accompany Medicaid reimbursements to hospitals to provide care for the poor to assist in the coverage of poor, uninsured patients – from a percentage formula to a block grant.38 An explicit purpose of such a conversion would be for the DSH program funding to be “reduced below current-law levels or its future growth limited to a slower rate than that at which Medicaid DSH payments would be made under current law, or both,” with states gaining flexibility is spending in return.39

These flexibility reforms would diminish the entitlement nature of Medicaid. As it currently exists, beneficiaries are entitled to the range of services defined in the state’s Medicaid plan as approved by CMS. The structural reforms tend to shift the program benefit from one assuring comprehensive medically necessary services to one designed by each state as that state, in that year, sees fit.

Medicaid’s Place in the Health Finance System
Medicaid is a sprawling program that has evolved over the last 40 years to fill many health finance gaps. Its central mission is to provide health coverage for people and services left out of both private health insurance and Medicare. As the absolute costs of health coverage has grown over time, the unorganized provision of services for those left out has become more and more difficult, and Medicaid – the formal financing system of last resort – has been used to provide predictable payment systems. It is, then, not one system but many.40 Most familiarly, it provides a comprehensive health insurance package for people who are poor and categorically eligible: children and their parents, people with disabilities and seniors. It is also the nation’s largest provider of long-term care coverage, both for seniors and people with disabilities. In addition, it finances a variety of services for people with mental retardation, developmental disabilities, and other disabilities.41 This expansive mission renders discussion of reform difficult. Simple prescriptions tend to be made by those limiting themselves to only a portion of the system (usually that providing basic health coverage for the non-disabled poor). Prescriptions that account for the complexity of the system tend to be, well, complex.

Medicaid could be thought to embody a positive response to an embraced social responsibility to ensure access to necessary health care for all, or as grudging recognition of market failure to be restrained and economized to the extent possible. Like a work by M. C. Escher, it is both, depending on the perspective of the viewer. In its earliest days, those administering Medicaid regarded their mission to be the financing of care of “high quality and in no way inferior to that enjoyed by the rest of the population,” through the elimination of a two-tier system of health care, assuring the poor access equal to that of the privately insured.42 Almost immediately, however, cost containment became a competing imperative in Medicaid regulation, creating a tension between service and retrenchment that continues to this day.43

This tension complicates today’s reform discussions. One reform perspective flows from a view that Medicaid should address well and with dignity the genuine health care needs of those otherwise unable to obtain appropriate care, with cost-containment consistent with appropriate access and quality. From this first perspective, Medicaid reform would mean improving access to services for beneficiaries, while ensuring that vendor overpayments and other inefficiencies were addressed. The competing perspective is much more grudging in its commitment to financing care for the poor. Although some services are acknowledged to be within government’s responsibility, the poor should be provided with basic services.
only, in order to lessen the burden on taxpayers and other sectors of state and federal budgets. From this second perspective, benefits should be limited to those essential to meeting the core mandates of the program in the most cost-effective manner possible.

Both perspectives acknowledge that financing care for the poor is a governmental function, although they differ on the richness of the care that should be funded. This dichotomy seems to leave out one perspective, which can be described as the “starve the beast” position. The “starve the beast” position, originally attributed to Ronald Reagan’s OMB director David Stockman, and now advanced by anti-tax advocates such as Grover Norquist, takes as a starting point that most social spending is wasteful and inappropriate.44 Acknowledging that many publicly funded social programs have a large constituency, it then advocates cutting taxes as a means to achieve surreptitiously what cannot be gained directly: at truly minimal government. Those advocating minimal government advocate services to the poor funded through voluntary charitable programs – harkening to the halcyon days of free clinics and pro bono services by physicians, allegedly choked by government interference.

Whether such a golden era ever existed is contested. It strains credulity, however, that a voluntary system would work in today’s expensive, complex, and bottom-line oriented medical care delivery systems. Its very implausibility explains why appeals for a return to charity are not made explicitly. Suggestions that charity could substitute for Medicaid should be examined for their relationship to cynical efforts to cripple government services without any realistic belief that a substitute source of coverage or care would arise. In less extreme forms, small government sentiments are clearly an important force in Medicaid reform efforts.

Medicaid and Private Insurance
Where does Medicaid fit into health finance today? Most obviously, it is the nation’s largest health care financer, spending approximately $300 billion in state and federal funds. It finances care for over 50 million people. It covers one-quarter of all American children, and one-third of births. It provides financing for care for millions of people with disabilities, as well as millions of seniors for whom Medicare simply fails to provide basic medical care.45 Medicaid’s beneficiaries are among the nation’s most expensive and vulnerable. It is the insurer of last resort for people with disabilities and they are an expensive group to cover. While people with disabilities make up only about 16 percent of Medicaid enrollees, they account for about 43 percent of expenditures.46 In addition, even those eligible for reasons other than disability are sicker and therefore more expensive than privately insured patients.47 Medicaid covers one in six Americans, and its enrollees are disproportionately disabled and/or sick.

Medicaid is not a program for an underclass disconnected from the labor economy. It is a program for the poor and permanently disabled, but it is also a program for low-income wage earners increasingly shut out of private insurance.

This description might suggest that Medicaid’s role in the nation’s health finance system is discrete, as it covers people excluded from private coverage. To an important extent that is true; many people covered by Medicaid will never be in the labor force, will never have resources to purchase coverage, and have health status that is in many contexts simply uninsurable. But Medicaid does interrelate with private health coverage in important ways that help to explain recent growth in its costs. First, Medicaid costs are counter-cyclical, as poor economic conditions leave a large number of formerly-insured workers unemployed, poor, and eligible for Medicaid. Economic slowdowns at the beginning of this decade, for example, left many workers (and their families), with no other source of health coverage but Medicaid.48 These spikes in Medicaid demand come, of course, at the same time that states are experiencing revenue shortfalls from the same economic downturns.

Medicaid is serving as a backup to the employment-based health insurance system in a long-term way as well. The 1990s was a time of increasing Medicaid coverage. The program severed its relationship with cash benefits programs, and expanded its reach into coverage of low income, categorically eligible workers and their families regardless of their connection to welfare.49 This expansion, intended to reduce the number of uninsured Americans, instead only picked up the slack left by a shrinking employment-based insurance system. During the late 1980s and early 1990s, the gap between high-wage and low-wage workers’ access to health insurance widened.50 In the last decade, private, employment-related coverage has continued to shrink, slowly during times of economic
expansion, and quickly during times of economic downturn.51 The decline in employment-based coverage is caused by increasing costs and the shift over time of premium costs from employers to employees.52 “New economy” service sector employers, an increasingly large sector of the employment market, are reluctant to take on the cost of coverage, leaving their employees uninsured or dependant on public assistance programs such as Medicaid.53

The explanation for recent increases in Medicaid costs is consistent with this description of the expansion of Medicaid’s mission and the weakening of the private, employment-based health insurance system. Few would disagree that Medicaid is an inexpensive system for health coverage on a per-person basis, in large part because of the tight reimbursement controls enforced by the states; the National Governors Association, for example, describes Medicaid as “extremely cost effective compared to private sector health care.”54 Medicaid costs have certainly grown, by perhaps one-third in the first three years of this decade. The increase is largely attributable to enrollment increases, however, as per-person cost increases were below those of private insurers.55 Medicaid’s enrollment has increased as states have sought to fill gaps in coverage, and as the private insurance system has decreased in relative strength. Medicaid is no longer a program for the poor with no connection to the labor economy. It is a program for the poor and permanently disabled, but it is also a program for low-income wage earners increasingly shut out of private insurance.

Reformers’ Vision of Medicaid

Medicaid, then, is a large program whose cost is growing rapidly due to enrollment increases rather than per-member cost inflation. It fills gaps left in our largely entrepreneurial health finance system, and covers very poor and severely disabled people with no chance of obtaining privately finance coverage. In addition, it is an increasingly important adjunct to the private employment-based health insurance system. This private system, long the backbone of American health finance, continues to deteriorate, leaving many low-income workers without access to private health insurance. As Medicaid’s consequently growing costs drive state and federal officials to examine the program for opportunities to cut expenses. Do the reform proposals signal a lessened commitment to Medicaid coverage for the poor and vulnerable? Some of these proposals, if carefully implemented, are simply good government tinkering with the program, and could improve it or at least reduce costs. Others nibble at the edges of the program, leaving the general structure in place while probably making the lives of the poor and disabled recipients marginally worse. The most significant proposals are the long-range proposals aimed at restructuring Medicaid to loosen states’ obligations to follow federal eligibility and benefits design requirements, and permit states to shift from a designed benefit toward a designed contribution model.

The proposals to tighten the payments to pharmaceutical firms raise few programmatic concerns, although the irony is thick with the Congress determined to avoid price regulation in Medicare Part D turning the screws in connection with Medicaid. The danger with reducing vendor or provider payments in Medicaid is that the rates will reach the point where the providers will withdraw their services. This has, of course, happened frequently in other Medicaid contexts, as states have set the reimbursement rates for dentists and physicians so low that beneficiaries have gone unserved.56 There is no suggestion that drug payment rates will reach that level. Furthermore, the political consequences of pharmaceutical companies withdrawing from Medicaid would likely be sufficiently serious as to give the firms pause.

Treatment of seniors’ access to nursing home care under some reform proposals raises interesting political concerns, but also significant programmatic concerns. The political issue is that seniors, including those well into the middle class, regard access to Medicaid reimbursement for nursing home care to be a well-entrenched entitlement, albeit with some need for fussing with “Medicaid planning” to adjust their assets and income. This group of seniors, unlike the poor families and children who are the target of other reductions, are politically active, and will object to the changes – as is evidenced by the AARP’s response to this proposed reform. The proposed asset transfer penalties will impoverish more seniors in need of nursing home care than do current rules. If aspects of the program must be cut, there is some logic to targeting asset transfer rules. This population is an “optional” category of eligibility, and one that is a disproportionate user of services. Although seniors are only 9 percent of beneficiaries, they account for 26 percent of Medicaid costs.57 They are financially better off than other Medicaid recipients, at least before the transfer of assets. They are well off, however, only when the comparison is with the most destitute. Compared to most, they live on a financial knife’s edge. Alternatives to reforms advocating more aggressive impoverishment are appropriate and available. Proposals made by the National Governors Association to create incentives, through tax credits, “long term care partnerships,” and other innovative methods offer the promise of Medicaid budget relief without the pain of impoverishment for seniors.58
The personal responsibility impulse behind many reform proposals seems poorly connected to experience with low income insureds. The basic notion behind proposals to increase cost-sharing is that sharing in the cost of care will cause the consumer to act wisely, parsimoniously, and – yes – responsibly. The trouble is that low-income patients, particularly those with chronic health conditions are poorly equipped to reduce their consumption wisely. That is, they do indeed reduce consumption, but they are just as likely to omit medically necessary care as care that could safely be refused. In addition, when the cost sharing comes not at the point of service, but rather in the form of premium co-insurance, a large percentage of participants will be forced to drop coverage, becoming uninsured, because low income beneficiaries have so many other calls on their income that the marginal increased cost of health insurance is simply unbearable. Personal responsibility rhetoric seems to match this situation quite poorly, and the cost sharing measures lead to the loss of medically necessary care and uninsurance as shrewd and prudent purchasing.

The remaining category of reform discussed above is the most important, the most varied, and the most likely to be seen as crucial to long-term reform: restructuring and flexibility. These reform proposals seek to devolve more programmatic authority to the states; seek to infuse “ownership society” concepts into Medicaid; and seek to further shift network formation and maintenance obligations in Medicaid from states to managed care plans, in part through partial shifts to defined contribution mechanisms. It is conceivable that a shift further away from Medicaid’s old way of doing business – fragmented fee for service care with little coordination or continuity of care – to a more context-sensitive and therapeutically appropriate system directed at wellness and disease management could result from these reform proposals.

Devolution could limit Medicaid’s countercyclical role. As is described above, Medicaid roles increase in bad economic times, as many people with private coverage lose their jobs and need health coverage for themselves and their families until they are able to reenter the labor force. Robust Medicaid expenditures are most needed when states are feeling least able to support them. Federal control over eligibility and benefits design provides some check on states’ reacting to bad economic times by cutting Medicaid programs just when they are needed the most. Recent events suggest caution; Missouri and Tennessee have reacted to tough economic times by cutting tens of thousands from Medicaid rolls, and other states, in less dramatic fashion, have responded in similar fashion. Continuing federal control over basic outlines of eligibility and benefits design provides some check on such activity.

Ownership society principles are likely to harm those with chronic illness. Some of the reform proposals, and some of the state waiver applications, are geared toward the “ownership society” movement’s preference for consumer control over some substantial portion of the funding for their care through “consumer-directed health care.” As is described above, the Medicaid population disproportionately comprises people with serious chronic illnesses. As I have described at length elsewhere, consumer-directed health care is likely to significantly disadvantage people with chronic illness, who have substantial, predictable, and medically necessary health care needs, and will therefore run through any personal spending account in short order. Where they require disease management and coordinated care, they are likely to be given autonomy and responsibility. Reform proposals suggesting personal savings accounts and other “ownership society” tools appear, in the Medicaid context to be a triumph of ideology over common sense.

Medicaid managed care imposes failed cost-saving principles on the poor. Managed care, as we knew it in the 1980s and 1990s, is dead. Managed care as it existed then was an experiment in outsourcing cost containment chores to entrepreneurial firms with little understanding of the mechanisms by which these firms would save money. This “black box” nature of managed care cost savings lead to abuses, consumer discontent, and ultimately to the wholesale rejection of the form by most consumers. The benefits realized through previous Medicaid managed care experiments have been mixed. Positive results seem to have flown only in states where the baseline prior to managed care was so low that almost any change would be likely to offer improvement. States saw little cost savings, and, except in those states with particularly poor access going into the managed care experiment, little improvement in patient care. The mistake in our reliance on tightly controlled managed care as a cost-savings device was that we did not anticipate cuts in services in return for cost savings; we expected the black box to produce free savings that it was unable to deliver. Managed care could, theoretically provide better coordination of care, particularly for those with chronic illness, than unorganized care could provide, but the improvement would not be free. It is unlikely that the proposed managed care-based reforms will achieve these advances any more than have similar reforms in the past. It seems unlikely that cost-containing reformers will select managed care models...
calculated to provide the sometimes-expensive coordinated care that chronically ill Medicaid beneficiaries need.

The reform proposals threaten the health care safety net. The fate of Medicaid beneficiaries and the uninsured are closely connected. For one thing, many people periodically move from one category to the other as they move in and out of the labor market. For another, the uninsured and people on Medicaid often rely on the same hospitals and other health providers for care. More to the point, however, the DSH program permits states to provide federally-supported funding to those facilities that serve Medicaid recipients and the uninsured. In addition, many providers of care to the uninsured rely on revenue received from Medicaid to bankroll care to the uninsured.69 The reduction in DSH payments will deprive safety net providers of funds necessary to treat the uninsured. Assignment of beneficiaries to cost-obsessed managed care organizations threatens to reduce the flow of Medicaid funding to those safety net providers.70

The savage irony of this aspect of the reform measures is that it promises to weaken the sources of care to the uninsured. To the extent the reforms represent a lessening of a commitment to Medicaid coverage for the poor – and the discussion above suggest that as a very real possibility – then the question must be asked, in favor of what alternative? It is nearly inescapable, given the deterioration of the private insurance market for low-income workers, that the alternative to alternative Medicaid coverage will be uninsurance. Those leaving Medicaid because they cannot meet cost-sharing requirements, because states choose to pare down their eligibility standards, or because the “defined benefit” put at their disposal is inadequate to provide full coverage, will be forced to rely on safety net providers for care. If, however, those same reforms cut funding to safety net providers, they will be out of luck.

Conclusion
Some of the Medicaid reform proposals now under discussion, including some likely to be adopted for short-term savings, will either not harm Medicaid beneficiaries or will “only” nibble at the edges of an already sparse system. Some proposals, however, including those likely to be pressed as long-term reforms, would dramatically change the shape of Medicaid, threatening substantial harm to its poor and disabled beneficiaries. In particular, those that would move Medicaid toward a defined contribution program; devolve enrollment and benefit design power to the states; and assign beneficiaries to entrepreneurial managed care organizations, threaten to limit care for beneficiaries and to force others from the program entirely. They therefore appear to be a lessening of commitment to Medicaid coverage for the poor and disabled, threatening to leave them uninsured. At the same time, reductions in DSH funding and reliance on managed care threaten to starve providers devoted to providing care for the uninsured. The long term reform proposals threaten to push beneficiaries from Medicaid coverage into the void, without a safety net.

References
5. See Pediatric Specialty Care, Inc. v. Arkansas Dept of Human Services, 364 F. 3d 925 (8th Cir. 2004); Rosie D. v. Swift, 310 F. 3d 230 (1st Cir. 2002); Arizican v. Odum, 290 F. 3d 176 (4th cir. 2002); Health Care for All, Inc. v. Romney, 2005 WL 1680677 (D. Mass 2005).
8. See Sanchez v. Johnson, 416 F. 3d 1051, 1060 (9th Cir. 2005).
on Medicaid and the Uninsured, State Specific Section 1115 Waiver Fact Sheets <http://www.kff.org/medicaid/waivers-index.cfm> (last visited November 1, 2005).


23. See R. Pear, supra note 18.


28. R. Pear, supra note 18, at 22 (quoting Representative Nathan Deal, Republican of Georgia).


31. See Medicaid Commission Report, supra note 7, at 12.

32. See National Governors Association: Short Run Medicaid Reforms, supra note 11, at 5.


34. Id. at 6-7.

35. 42 C.F.R. 440.230(b).


38. See Congressional Budget Office, supra note 37, at 15-16.

39. Congressional Budget Office, supra note 37, at 15.


41. See Rosenbaum, id. at 636.

61. See B. J. Wright, et al., *supra* at 1112-1113.


67. Id.


70. See T. Stoltzfus Jost, *supra* note 66 at 128.
Advances in science and technology frequently raise new ethical, legal, and social issues, and developments in neuroscience and neuroimaging technology are no exception. Within the field of neuroethics, leading scientists, ethicists, and humanists are exploring the implications of efforts to image, study, treat, and enhance the human brain.1

This article focuses on one aspect of neuroethics: the confidentiality and privacy implications of advances in functional magnetic resonance imaging (“fMRI”). Following a brief orientation to fMRI and an overview of some of its current and proposed uses, this article highlights key confidentiality and privacy issues raised by fMRI in the contexts of health care, research, employment, insurance, criminal justice, litigation, and cognitive privacy.

Functional Magnetic Resonance Imaging

Magnetic resonance imaging (“MRI”) uses radiofrequency waves and a strong magnetic field to provide detailed images of internal organs and tissues. Functional MRI is based on the same technology as MRI. However, instead of imaging static, soft-tissue structures, fMRI measures localized changes in the brain that occur when an individual performs a mental task, such as viewing an image, responding to a question, or listening to a voice. By subtracting a control image from an experimental image, physicians and scientists can create maps showing the regions of the brain to which a surplus of oxygenated blood flows when an individual performs a task.2

Functional MRI has several clinical applications. Preoperatively, physicians and scientists use fMRI to study patients while they complete a battery of mental tasks and to identify the regions of the brain that are associated with tactile, motor, language, and visual functions.3 Scientists believe these cortical maps can help neurosurgeons assess surgical risk, plan surgical routes, and direct intraoperative electrophysiological procedures,4 although others caution against the leap from functional imaging to functional neurosurgery.5

Scientists also use fMRI to examine the physiological correlates of well-known social psychological phenomena. For example, several groups of scientists have used fMRI to explore the neural substrates involved when research subjects view faces of white and black individuals.6 In one study, the authors concluded that representations of social groups that differ in race evoke differential amygdala activity related to unconscious social evaluation.7 Although the authors stated that their results cannot be taken as a means of testing for racism in individuals, others speculate that refinement of fMRI technology could unveil racial preferences and prejudices.8

The neural correlates of deception also have been studied using fMRI.9 In one popular study, scientists asked subjects to hold a 5 of Clubs playing card in their pocket and to deny that they held the card while their brains were being imaged.10 The scientists concluded that cognitive differences between deception and truth have neural correlates detectable by fMRI and that refinements in study design could establish an activation pattern predictive of deception on an individual level.11 Although some scientists and ethicists caution against premature use of fMRI to detect deception in non-research settings,12

Stacey A. Tovino

Stacey Tovino is Visiting Assistant Professor at the University of Houston Law Center Health Law & Policy Institute and a Ph.D. candidate at the University of Texas Medical Branch Institute for the Medical Humanities. Her dissertation (“The Visible Brain: Confidentiality and Privacy Implications of Functional Magnetic Resonance Imaging”) explores the issues raised in this article in more detail.
other individuals speculate that fMRI might enable government and criminal justice officials to determine whether criminal suspects and terrorists are lying or telling the truth.13

Scientists also have used fMRI to study the neural correlates of altruism and social cooperation.14 In one study involving two separate experiments, scientists scanned the brains of thirty-six women as they played the Prisoner’s Dilemma, a game in which two players independently choose whether to cooperate with each other or betray each other for immediate gain.15 The scientists concluded that mutual cooperation was associated with consistent activation in regions of the brain linked to reward processing.16 Others speculate that employers might want to use fMRI to recruit applicants for employment who experience more or less pleasure from cooperation, depending on the requirements of the job.17

Functional MRI has been used to study a range of other conditions and characteristics including, but certainly not limited to, Alzheimer’s disease,18 major depression,19 schizophrenia,20 bipolar disorder,21 dyslexia and hyperlexia,22 minimal consciousness,23 pedophilia,24 cocaine addiction,25 compulsive gambling,26 satiety and obesity,27 extraversion,28 self-consciousness29 maternal and romantic love,30 and sexual arousal,31 as well as individuals’ preferences regarding soft drinks,32 automobiles,33 campaign advertisements,34 and the content of movie trailers.35

Confidentiality Implications
An oft-stated principle is that physicians and scientists have an ethical and legal duty to maintain the confidentiality of medical and study records in their possession. Applying existing confidentiality obligations to neuroimages and related data sets, reports, and interpretations (collectively, “neuroimaging information”) raises several questions. First, do existing confidentiality protections apply to the growing number of scientists who create, use, and disclose neuroimaging information? Second, do existing protections apply to neuroimaging information that is interpreted to reveal social qualities and characteristics, but not particular physical or mental conditions? Third, when is neuroimaging information individually identifiable? Finally, how can scientists disclose neuroimaging information to neuroimaging databanks in accordance with publication requirements and peer suggestion without violating confidentiality requirements?

The Department of Health and Human Services’ standards for the privacy of individually identifiable health information (the “Privacy Rule”)36 will apply to some, but not all, of the individuals and institutions that create, maintain, or desire to obtain neuroimaging information. Very generally, the Privacy Rule regulates covered entities’ use and disclosure of protected health information.37 Covered entities include health care providers who transmit health information in electronic form in connection with certain standard transactions, as well as health plans and health care clearinghouses.38

Many health care providers (including hospitals, imaging centers, radiologists, and neurosurgeons) transmit health information in electronic form in connection with claims sent to health insurers for reimbursement and other standard transactions. These health care providers, as well as the health insurance companies to which the providers’ claims are sent, constitute covered entities and must maintain the confidentiality of their protected health information in accordance with the Privacy Rule or risk civil and criminal penalties.39 However, the Privacy Rule generally does not apply to many other individuals and institutions reported to have an interest in neuroimaging information, including federal and state governments, criminal justice officials, employers, life insurance companies, litigants, and marketing companies.

Understanding how the Privacy Rule applies to the growing number of scientists who use fMRI technology can be confusing. The Privacy Rule generally does not apply to: (1) scientists who do not provide health care as part of their research; and (2) scientists who do provide health care but who do not electronically bill for such health care or do not otherwise transmit health information in electronic form in connection with a standard transaction.40 However, the Privacy Rule does apply to the scientists described in the preceding sentence if they are employees or workforce members of a university that: (1) is a single legal entity; (2) has a

Do existing confidentiality protections apply to the growing number of scientists who create, use, and disclose neuroimaging information?
is taken for the purpose of studying one-time deception that does not constitute pathological lying (“I do not have the 5 of Clubs card”)? What about a neuroimage that shows amygdala activity interpreted as unconscious social evaluation? What about a neuroimage that is interpreted to reveal an individual’s preference for a particular soft drink, automobile, campaign advertisement, or movie trailer? A very technical argument exists that these latter pieces of neuroimaging information do not constitute health information under the Privacy Rule because they do not relate to the physical or mental health or condition of an individual or the provision of health care to an individual.

A third issue is whether neuroimaging information is individually identifiable. The Privacy Rule only protects health information if it is individually identifiable, meaning that the information identifies the individual who is its subject or there is a reasonable basis to believe that the information could be used to identify the individual.44 Neuroimaging information that contains a patient’s name, telephone number, social security number, medical record number, account number, or other similar identifier generally is considered individually identifiable.45

When is neuroimaging information sufficiently de-identified such that a covered entity can use and disclose the information without regulation by the Privacy Rule? The Privacy Rule contains a de-identification safe harbor that requires removal of eighteen different types of identifiers before data is considered de-identified.46 Satisfying the de-identification safe harbor in the context of neuroimaging is slightly more difficult than in the traditional medical record context because of reports that individuals’ facial features can be reconstructed from high resolution fMRIs.47 Because the Privacy Rule’s safe harbor requires removal of full-face photographs and comparable images, covered entities also must remove from neuroimaging information any data or voxels from which an individual’s face can be recognized or reconstructed.

Because the Privacy Rule’s safe harbor requires removal of full-face photographs and comparable images, covered entities also must remove from neuroimaging information any data or voxels from which an individual’s face can be recognized or reconstructed.

As a condition of publication, the Journal of Cognitive Neuroscience requires its authors to submit their complete fMRI study data to the fMRI Data Center, a neuroimaging databank located in Hanover, New Hampshire.48 At least two other journals strongly encourage submission of complete study data to neuroimaging databanks to support the findings, outcomes, and claims in the journals’ published articles, and some scientists encourage the disclosure of neuroimaging information to neuroimaging databanks to speed the understanding of cognitive processes and the neural substrates that underlie them.49 Thus, a fourth issue is how covered entities can disclose neuroimaging information to neuroimaging databanks without violating the Privacy Rule. The Privacy Rule allows covered entities to disclose de-identified information to databanks without the prior written authorization of the individuals who are the subjects of the information.50 Accordingly, if a covered entity removes from the data it sends to a databank all eighteen identifiers listed in the Privacy Rule’s de-identification safe harbor, including any data from which a subject’s face can be recognized or reconstructed, the Privacy Rule would allow the disclosure to the databank.51

Several other laws establish confidentiality protections that may apply to individuals who create, maintain, or desire to obtain neuroimaging information. For example, federal protection of human subjects regulations (the “Common Rule”) requires institutional review boards approving federally-funded research to ensure that, “[w]hen appropriate, there are adequate provisions to...protect the confidentiality of [subjects’] data.”52 The Office for Human Research Protections has stated that the “adequate provisions” language in the Common Rule requires investigators to replace names and other identifiers with codes and to store paper and electronic research records securely.53 Unlike the Privacy Rule, the Common Rule protects all federally-funded, human subject study data,54 not just health information in the possession of covered entities. Similar to the Privacy Rule, the Common Rule contains a de-identification standard that would require the removal of data from which the identity of the subject can or may be readily ascertained, including data or voxels from which an individual’s face can be recognized or reconstructed.55

In addition to federal confidentiality protections, many states have medical practice acts, hospital licensing laws, imaging center licensing laws, and other similar laws and regulations that require certain individuals and institutions to maintain the confidentiality of health information in their possession. These laws typically define health information as information that relates to the diagnosis, treatment, or prognosis of patients.56 However, many of these laws do not: (1) extend their protections to scientists who do not provide health care to patients; (2) state whether their protections extend to social information in addition to health information; or (3) specify whether, or the conditions under which, a provider may disclose neuroimaging information to a neuroimaging databank.
In summary, advances in neuroimaging raise a number of confidentiality issues. Which existing confidentiality laws protect social information created and maintained by non-provider scientists? Does the Common Rule adequately protect the confidentiality of neuroimaging information in the possession of scientists? Should Congress or state legislatures enact new laws providing heightened confidentiality protections for neuroimaging information? If so, how would such laws define the neuroimaging information to be protected? How would we reconcile such heightened protections with other laws that allow the disclosure of confidential information without authorization for certain public policy purposes? Would heightened confidentiality protections for neuroimaging information suggest that all neuroimaging information is sensitive or stigmatizing, even though it may not be? Finally, what lessons can we learn from federal and state efforts to establish heightened confidentiality protections for genetic information?

Privacy Implications

Patients voluntarily disclose some information to health care providers to obtain health care, and human subjects consent to scientists’ obtaining some personal information during research studies. But, what if a provider or scientist discovers a condition that the patient or subject would have preferred to keep private? Recent studies analyze the extent to which scientists have discovered arteriovenous malformations, brain tumors, developmental abnormalities, and other conditions in healthy controls who participate in neuroimaging research. The question becomes, can fMRI violate an individual's interest in keeping certain information private? Physicians and scientists might be able to minimize an individual's perception that her privacy has been violated by identifying, as part of the informed consent process, the type of information potentially discoverable by fMRI, and by negotiating a notification and treatment referral process to be followed in the event of an unanticipated finding.

Although the cost to employers of conducting their own fMRI tests likely would be prohibitive, employers might be interested in obtaining the results of past fMRI tests to assist in decision making relating to the selection and retention of employees. Thus, the issue in the employment context is whether applicants for employment and employees have the ability to keep their neuroimaging information private or whether they can be forced to reveal such information pursuant to a compelled authorization. A handful of laws regulate employers’ use of employment tests, medical examinations, and related inquiries. Two examples include Title I of the Americans with Disabilities Act (the “ADA”) and the Employee Polygraph Protection Act.

Among other activities, Title I of the ADA regulates certain employers’ use of qualification standards, employment tests and other selection criteria that screen out or tend to screen out individuals with disabilities on the basis of such disabilities (the “screening provisions”). Equal Employment Opportunity Commission regulations interpreting Title I define disability to include physical and mental impairments (including neurological disorders, mental illnesses, and specific learning disabilities) that substantially limit one or more major life activities of an individual. EEOC regulations also clarify, however, that pedophilia, compulsive gambling, homosexuality, and certain other characteristics do not constitute disabilities protected by the ADA. The result is that the ADA’s screening provisions would regulate a covered employer’s use of fMRI in an attempt to screen out individuals who have depression, schizophrenia, or bipolar disorder if such conditions substantially limit a major life activity of the individual tested. On the other hand, the screening provisions would not regulate attempts to screen out individuals based on fMRI “findings” of pedophilia, compulsive gambling, and homosexuality.

Title I of the ADA also regulates the conduct and timing of medical examinations and inquiries. The EEOC has issued enforcement guidance that defines a medical examination as a procedure or test that seeks information about an individual’s health or physical or mental impairments. Although a number of factors are relevant in determining whether a procedure or test is a medical examination, the EEOC clarifies that the term includes tests that provide evidence leading to the identification of conditions listed in the American Psychiatric Association’s most recent Diagnostic and Statistical Manual of Mental Disorders, including anxiety, depression, and certain compulsive disorders that have been studied by fMRI. The EEOC also clarifies, however, that psychological tests designed and used only to measure honesty, tastes, and habits are not medical examinations. A determination of how the ADA’s medical examination and inquiry provisions apply to particular fMRI tests will require application of the factors and interpretations set forth by the EEOC in its enforcement guidance.

With some exceptions, the federal Employee Polygraph Protection Act ("EPPA") prohibits employers from requiring employees to submit to lie-detector tests, defined to include polygraphs, deceptrons, voice stress analyzers, psychological stress evaluators, and "any other similar device...that is used, or the results of which are used, for the purpose of rendering a diagnostic opinion regarding the honesty or dishonesty of an individual." Thus, the EPPA could be interpreted to prohibit employers from requiring neuroimaging examinations that could form the basis of an opinion regarding an individual's dishonesty.

Functional MRI has privacy implications beyond the health care, research, and employment contexts. In the insurance context, the concern is that health, life, and auto insurers will use individuals’ neuroimaging information in an attempt to predict future illness, a propensity to violence, or other conditions or charac-
teristics relevant to underlying deci-
sions. Is special legislation or reg-
ulation needed to prevent the gather-
ing or use of neuroimaging information by insurance compa-
nies? If so, how can recent legisla-
tive efforts to restrict the use of ge-
etic information by health and life
insurers guide efforts to protect the privacy of neuroimaging informa-
tion? In the criminal justice and civil lit-
igation contexts, the issue is whether fMRI has the potential to violate the
privacy rights of individuals sus-
pected of being terrorists, individuals
suspected of engaging in other crim-
nal activity, and participants in civil
litigation. For example, does an fMRI of a criminal suspect’s brain constitute a search under the Fourth Amendment? Could an fMRI of a criminal suspect’s brain violate the Fifth Amendment privilege against compulsory self-incrimination? Can counsel use fMRI during voir dire to exclude jurors whose neuroimages are interpreted to re-
veal racial prejudices? Finally, and regardless of the context, do all indi-
viduals have the right to cognitive privacy, or the privacy of their own
thoughts? Can state action that punishes an individual or holds an
individual responsible for thoughts, but not actions, violate the individ-
ual’s cognitive privacy?

Although neuroimaging research has improved our understanding of the neural bases of personality, behavior, and consciousness, it raises additional questions relating to the concept of the self, emotional and
moral judgment, prediction of future illness, unanticipated findings, truth
telling, social and legal responsibility and, central to this article, confiden-
tiality and privacy. Existing confiden-
tiality and privacy protections tend to be tied to the concepts of physical and mental health, not indi-
viduals’ thoughts, preferences, and social conduct. Identification of po-
tential gaps in coverage can inform policy discussions about the need to
protect social qualities and charac-
teristics and to promote mental au-
tonomy as attempts to transfer fMRI
technology beyond the research con-
text are made.

References
science and the Law: Brain, Mind, and the Scales of Justice (New York: Dana
Press, 2004); S. J. Marcus, ed., Neu-
roethics: Mapping the Field (New York: Dana Press, 2002).
2. J. Illes and E. Racine, "Imaging or Imag-
nining? A Neuroethics Challenge In-
tional Magnetic Resonance Imaging Pro-
cedure for Preoperative Mapping of Cor-
tical Areas Associated with Tactile, Motor, Language, and Visual Functions,”
4. Id. at 718-720.
6. See, e.g., M. T. Wheeler and S. T. Fiske,
“Controlling Racial Prejudice and Stere-
typing: Social-Cognitive Goals Affect Amygdala and Stereotype Activation,”
tive Neuroscience 12, no. 5 (2000): 729-
38; A. J. Hart, et al., “Differential Re-
spose in the Human Amygdala to Racial Outgroup vs. Ingroup Face Stimuli,” Neu-
7. See Phelps, et al., supra note 6, at 714.
8. H. Breithaupt and K. Weigmann, “Ma-
nipulating Your Mind: What Will Science Discover About Our Brains, and How Are We Going to Deal with It?” EMBO Re-
ioral Neuroscience 118, no. 4 (2004): 852-
56; F. A. Kozel, et al., “A Pilot Study of Functional MRI Brain Correlates of De-
ception in Healthy Young Men,” Journal of Neuropsychiatry and the Clinical Sci-
cences 16, no. 3 (2004): 295-305; G. Gani, et al., “Neural Correlates of Differ-
tent Types of Deception: An fMRI In-
vestigation,” Cerebral Cortex 13, no. 8
nance Imaging,” Human Brain Mapping, 15, no. 4 (2002): 157-64; D. D. Lan-
gleben, et al., “Brain Activity during Sim-
ulated Deception: An Event-Related Functional Magnetic Resonance Study,”
NeuroImage 15, no. 3 (2002): 727-32; S.
A. Spence, et al., “Behavioural and Func-
tional Anatomical Correlates of Decep-
tion in Humans,” NeuroReport 12, no. 13
11. Id. at 731.
13. B. W. Orenstein, “Guilty? Investigating fMRI’s Future as a Lie Detector,” Radi-
ology Today 6, no. 10 (2005): 30.
ward Pathways,” NeuroReport 15, no. 16
16. Id. at 403.
18. See, e.g., A. J. Golby, et al., “Memory En-
coding in Alzheimer’s Disease: An fMRI Study of Explicit and Implicit Memory,”
tive Disorder,” British Journal of Psy-
22. See, e.g., S. Ruff, et al., “Neural Sub-
strates of Impaired Categorical Percep-
23. N. D. Schiff, et al., “fMRI Reveals Large-
Scale Network Activation in Minimally
Conscious Patients,” Neurology 64, no. 3
848
THE JOURNAL OF LAW, MEDICINE & ETHICS CURRENTS IN CONTEMPORARY ETHICS


45. See, e.g., id.


49. See J. Tierney, “Using M.R.I.’s to See that the Information is not Individually Identifiable to Determine that the Information is not Individually Identifiable,” Id. § 164.514(a).

50. Id. § 164.514(b)(1). A second method of de-identifying information allows a person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods for rendering information not individually identifiable to determine that the information is not individually identifiable. Id. § 164.514(b)(1).


52. Journal of Cognitive Neuroscience, “Submission Guidelines,” at http://mitpress.mit.edu/catalog/item/default.asp?type=4&tid=12&xid=6&xcid=0 (“All papers accepted in JOCN will be required to submit their fMRI data to the FMRI Data Center.”) (last visited September 15, 2005).


54. C.F.R. § 164.514(b)(2)(i). If one or more of the eighteen identifiers remain in the information, the Privacy Rule would require the covered entity to: (1) obtain a prior written authorization containing certain required elements and statements from each subject whose identifiers are disclosed to the database; or (2) satisfy a research-related exception to the Privacy Rule’s authorization requirement. One potentially applicable exception permits an institutional review board or privacy board to waive the authorization required by the Privacy Rule if certain criteria are satisfied. See id. § 164.508(c)(1) and (2); id. § 164.512(i).

55. Id. § 46.111(a)(7).


57. See 45 C.F.R. § 46.102(f) (2005) (protection data that results from “intervention or interaction with [an] individual” or “identifiable private information”).

58. Id. § 46.102(f).

59. See, e.g., CAL. CIV. CODE §§ 56.05-37, 56.05(g) (West 2005) (California Confidentiality of Medical Information Act provisions protecting “medical information,” defined as “individually identifiable information, in electronic or physical form, in possession of or derived from a provider of health care, health care service plan, pharmaceutical company, or contractor regarding a patient’s medical history, mental or physical condition, or treatment”); TEX. HEALTH & SAFETY CODE ANN. §§ 241.151-153, 241.151(2) (Vernon 2005) (Texas Hospital Licensing Law provision establishing confidentiality protections for “health care information,” defined as information that relates to the “history, diagnosis, treatment, or prognosis of a patient”).


61. Id. § 144.


67. Id. § 1630.2(g).

68. Id. § 1630.3(d)(1) (pedophilia not disability); id. § 1630.3(d)(2) (compulsive gambling not disability); id. § 1630.3(e) (homosexuality not impairment so not disability).


71. Id.

72. Id.


74. Id. § 2001(3).


76. Greely, supra note 57, at 124-25.


80. Greely, supra note 57, at 146.

81. Id. at 136.


83. See, e.g., Fallo v. Connecticut, 302 U.S. 319, 326-27 (1937) (“[F]reedom [of thought]...is the matrix, the indispensable condition, of nearly every other form of freedom”); Jones v. Opelika, 316 U.S. 584, 618 (1942) (“[F]reedom to think is absolute of its own nature; the most tyrannical government is powerless to control the inward workings of the mind”); Abood v. Detroit Board of Education, 431 U.S. 209, 234-35 (1977) (“[A]t the heart of the First Amendment is the notion that an individual should be free to believe as he will...”); Ashcroft v. Free Speech Coalition, 535 U.S. 234, 253 (2002) (“The right to think is the beginning of freedom... ”). See generally Brief of Amici Curiae Center for Cognitive Liberty & Ethics, at 3-7, Sell v. United States, 539 U.S. 166 (2003) (No. 02-5664) (arguing that the First Amendment guarantees freedom of thought).
Incompetent Decisionmakers and Withdrawal of Life-Sustaining Treatment: A Case Study

Lance Lightfoot

One of the most challenging and rewarding roles for in-house hospital attorneys is serving as a member of their hospital’s Bioethics Committee (the “Committee”). As a member of the Committee, an attorney assists in developing institutional ethics policies and guidelines, and also participates in ethics consultations involving disputes about patient care. Institutions such as the Author’s employer, Texas Children’s Hospital, promote open and honest communications between members of a patient’s health care team and the patient’s parents and family; however, when communications break down, the Committee’s goal is to provide an objective forum where disputes can be discussed and hopefully resolved in a professional, ethical manner.

The vast majority of ethics consultations at Texas Children’s result in some form of compromise and resolution between the patient’s family and the patient’s physician(s). One consultation, however, did not follow this trend. The consultation involved a baby boy named Sun who had been born with a lethal genetic condition. Sun’s physicians at Texas Children’s believed withdrawing care, as provided for under the Texas Advance Directives Act (“the Act”), and allowing Sun to die naturally was medically appropriate and the most ethical course of treatment for the tragic situation. In their view, to continue providing care to Sun would only extend his suffering. Wanda Hudson, Sun’s mother, refused to consent to the withdrawal of care, however, so the health care team asked the Committee to intervene.

While all ethics consultations are difficult, the Hudson ethics consultation was particularly complicated due to Ms. Hudson’s peculiar comments and behavior before, during and after the consultation. Ms. Hudson’s behavior caused the Committee to have concerns that she – Sun’s legal guardian and sole decision-maker – had some type of psychological impairment, and the Committee struggled with whether and how her condition should affect the Committee’s deliberations and decision. After two ethics consultations regarding the matter, the Committee ultimately decided to support Sun’s physicians’ plans to withdraw care. However, based on the Committee’s concerns about Ms. Hudson’s competence, the Committee and hospital took the unusual step of assisting Ms. Hudson with retaining an attorney and agreeing to pay her legal bills. After retaining an attorney, Ms. Hudson filed a lawsuit against Texas Children’s to prevent the physicians from carrying out the withdrawal of care.

The case, one of the first in which a court ultimately was asked to approve the withdrawal of life-sustaining treatment for a child over a parent’s objections, drew national media attention and triggered debate over the propriety of giving physicians and hospitals so much power in end-of-life cases. While that issue can be debated elsewhere, this article will focus on a unique but important issue raised by the case: if the person solely responsible for making health care decisions for a minor patient with a terminal or irreversible condition is believed to be mentally incompetent, how should the child’s health care providers proceed with withdrawing care if such withdrawal is deemed to be medically appropriate? Moreover, how should the ethical health lawyer render advice that is both legally accurate and ethically appropriate in this difficult situation?
This article will address these questions by using the Hudson ethics consultation as a case study. The article will first review the circumstances of Sun’s birth and his medical condition, as well as the concerns Sun’s physicians and the hospital staff had about Ms. Hudson’s mental status. Next, the article will address the manner in which the Committee became involved in the case, and the effect Ms. Hudson’s mental status had on the Committee’s deliberations. An overview of the Texas Advance Directives Act, with specific focus on the statutory process for withdrawing care, is helpful in providing context within which the Committee was operating. In the Author’s view, the unique fact pattern of the Hudson case exposed a need to expressly require persons responsible for making decisions on behalf of such patients to be competent. The article will briefly describe the various efforts made by the Committee and the hospital to resolve the matter in a way that took into account Sun’s best interests while also addressing Ms. Hudson’s condition – as well as why, after all other efforts failed, the Committee and the hospital decided to assist Ms. Hudson with retaining an attorney to represent her and agreed to pay her legal bills, recognizing it was quite likely Ms. Hudson would bring a legal claim to oppose the Committee’s and the hospital’s decision. The article will conclude by proposing several changes to the advance directives and guardianship laws that could assist hospitals, physicians, patients, and their families who may be faced with situations similar to Sun’s in the future.

Ms. Hudson and Sun
Wanda Hudson gave birth to Sun on September 25, 2004, at St. Luke’s Episcopal Hospital in Houston, Texas. She had received no prenatal care prior to giving birth, so she was unaware that her baby would be born with significant physical and mental disabilities. When Sun was born he had dwarf-like features, with very short appendages and an enlarged head, and he experienced significant respiratory distress from the moment he was delivered. He was immediately transferred from St. Luke’s to Texas Children’s, where he was admitted to the hospital’s Level III neonatal intensive care unit (“NICU”), the largest such unit in the nation. He was placed on a respirator and a feeding tube was inserted.

Because there are numerous types of neonatal dwarfism, Sun’s physicians at Texas Children’s ordered genetic testing to determine the specific diagnosis. An outside lab conducted the testing and came back with dire results: Sun was born with thanatophoric dysplasia, which is a rare, fatal condition. While thanatophoric dysplasia causes severe mental and physical disabilities, what ultimately contributes to death is an abnormally narrow chest cavity that restricts the baby’s ability to breathe. Further, the chest cavity does not grow normally, so the baby’s breathing problems only worsen until the baby slowly suffocates to death.

Meanwhile, after giving birth to Sun at St. Luke’s, Ms. Hudson was involuntarily committed to a Houston psychiatric hospital based in part on her insistence to St. Luke’s physicians and staff that Sun was not from this earth and that he was the son of the sun in the sky. It appears that after evaluation Ms. Hudson was deemed not to be a threat to herself or others, so she was released by the facility and came to visit Sun at Texas Children’s.

The physicians and staff in the NICU had been somewhat prepared for Ms. Hudson’s arrival based on notes from St. Luke’s staff in Sun’s medical records and discussions between St. Luke’s and Texas Children’s staff concerning her behavior after Sun’s birth. When she arrived, Sun’s physicians informed Ms. Hudson of Sun’s condition and expressed their sympathy to her. Members of the hospital’s social work department and clergy also spoke with Ms. Hudson in an effort to assist her in coping with Sun’s impending death. Ms. Hudson, however, did not respond well to any of these discussions, repeatedly stating Sun was a normal baby who just needed time to develop.

Sun’s physicians decided to give Ms. Hudson a few days to appreciate the gravity of the situation, with the expectation that she would eventually agree with the recommendation to withdraw care. However, when it became clear to Sun’s physicians that Ms. Hudson was not inclined ever to agree with them, they asked the Committee to become involved.

When it became clear to Sun’s physicians that Ms. Hudson was not inclined ever to agree with them, they asked the Committee to become involved.

The Bioethics Committee and the Advance Directives Act
When the Committee at Texas Children’s is asked to consult on an end-of-life case where the patient’s physicians and family disagree, the attorney’s primary concern is to ensure the Committee’s actions comply with the Texas Advance Directives Act (the “Act”), specifically the provisions of Section 166.046 of the Act. Section 166.046 sets forth the process under which physicians and hospitals may, under certain circumstances, withdraw life-sustaining treatment from a patient, even against the wishes of the patient’s family.

In short, the statutory provisions require that a physician’s refusal to honor a treatment decision made by or on behalf of a patient be reviewed by a hospital ethics or medical committee. At Texas Children’s, the Committee specifically is charged with performing this function. The person responsible for making treatment decisions for the patient must be provided with 48 hours advance notice of the review process, and
given an opportunity to attend the meeting. Once a decision is made regarding the patient’s care, the Committee is required to provide written notification of the decision to the person responsible for making decisions for the patient. If the Committee agrees with the physician’s decision to withdraw life-sustaining treatment from the patient, the physician is required to make a reasonable effort to transfer the patient to a physician or facility willing to continue to care for the patient. The patient must be provided available life-sustaining treatment pending the transfer, but the physician and hospital are not obligated to provide such treatment after the tenth day after the written decision of the Committee is provided to the person responsible for making health care decisions for the patient unless ordered to do so by an appropriate district or county court. If the treating physicians, other health care providers and the hospital follow the process outlined in the Act, they are immune from civil liability for withdrawing life-sustaining treatment from the patient; moreover, the physicians and other health care providers are not subject to any criminal liability or disciplinary action by their respective licensing boards, unless they failed to exercise reasonable care.

Prior to the Hudson matter, the Committee at Texas Children’s had been very successful in resolving disagreements between physicians and their patients’ families. That would not be the case after Ms. Hudson attended the ethics consult to discuss Sun’s prognosis and treatment. The full Committee first met to discuss Sun’s case on October 13, 2004. Ms. Hudson and Sun’s physicians attended the meeting to provide the Committee with their positions on Sun’s diagnosis, prognosis and how Sun’s care should proceed. Ms. Hudson reiterated her belief that Sun was normal and would grow if given the opportunity, and spoke in a rambling, defensive and sometimes incoherent manner. Sun’s physicians provided testimony and medical literature supporting their belief that providing further care to Sun was futile and medically inappropriate. After discussing the matter, the Committee agreed with Sun’s physicians that withdrawing care was medically appropriate, but were concerned that if they sent the letter to Ms. Hudson explaining the Committee’s decision and notifying her that care would be withdrawn in ten days, she would not understand the gravity of the situation nor the consequences of inaction.

One could argue that this concern was irrelevant, since Sun’s best interests should have been paramount and Ms. Hudson’s understanding of the process was secondary. Indeed, because Ms. Hudson, as Sun’s mother, had the legal right to make health care decisions on behalf of Sun under the Act and the Act does not explicitly require decision-makers to be competent — legally, the Committee most likely could have gone forward at this point and started the process for eventual withdrawal of care. From an ethical standpoint, however, the Committee did not want to risk the possibility of withdrawing care from Sun and discovering afterwards that Ms. Hudson did not fully understand what was going on. Therefore, no letter was sent to Ms. Hudson and the Committee decided to delay its decision until it determined whether any other options were available.

Options Explored by the Committee
The Committee asked legal counsel to research the availability of alternate means to resolve the matter. The Committee specifically requested options that would serve to mitigate the impact of Ms. Hudson’s perceived mental incompetence on the Committee’s deliberations concerning Sun’s treatment. Legal counsel explored several options, focusing first on whether Sun’s father or any of Ms. Hudson’s relatives would agree to become involved in the decision-making process. When that option failed, legal counsel examined other less obvious means to remove Ms. Hudson from the decision-making role.

Other Family Members
One of the first options explored by the Committee was locating Sun’s father or another family member to assist the Committee in communicating with Ms. Hudson, and to determine his/her willingness to take steps to intervene in the decision-making process. The Committee was unsuccessful on both fronts. Sun’s father never came forward and could not be identified — Ms. Hudson repeatedly told Sun’s physicians that Sun’s father was the sun in the sky. Further, at the time she gave birth to Sun Ms. Hudson was living with her parents, who over time had several discussions with Texas Children’s staff about Sun’s care and the Committee process. They were unwilling, however, to try to influence Ms. Hudson’s decisions or otherwise get involved in the matter, as were several other members of Ms. Hudson’s family.

Child Protective Services
Under applicable Texas law, physicians and hospital personnel are required to contact Child Protective Services ("CPS") if they suspect a child is being, or has been, abused or neglected. The Committee explored the possibility of reporting Sun’s case to CPS based on the argument that Ms. Hudson’s decision to maintain Sun on life support could be considered abuse, in that it served no medical purpose and prolonged his suffering. When contacted regarding Sun, however, CPS officials refused to become involved, taking the position that Sun was being provided with the best medical care available and Ms. Hudson was not actively harming her child.

Adult Protective Services
The Committee also explored whether Adult Protective Services ("APS") would become involved in the case, based on the Committee’s concern that Ms. Hudson was suffering from some form of psychological impairment that was affecting her decision-making ability. The Committee learned, however, that APS focuses its efforts on investigating reports of abuse, neglect and/or
exploitation of adults,20 and did not consider Ms. Hudson’s case to be within the agency’s jurisdiction.

Hospital Guardianship
The Committee also explored the possibility of the hospital pursuing guardianship or being appointed as guardian ad litem for Sun. Under the Texas Family Code, however, only certain persons have standing to file a suit affecting the parent-child relationship, and a hospital is not one of those persons.21 The hospital’s only chance of securing a guardian for Sun was through CPS, which has the authority under the Family Code to investigate claims of abuse or neglect and to take possession of a child.22 This was not a viable option, however, since CPS officials had already informed the Committee that they would not intervene in the matter. The Committee seemed to have run out of options.

The Committee’s Decision and the Resulting Lawsuit
The Committee reconvened on November 18, 2004. Ms. Hudson was again provided notice of the meeting and invited to attend. The Committee recommended to Ms. Hudson that she bring an attorney to the meeting to assist her in understanding the process, but she attended alone. Sun’s physicians attended as well, but neither they nor Ms. Hudson provided the Committee with much new information. Of course, Sun was now almost two months old, and his medical condition had only worsened since the last Committee meeting.

After hearing from both sides, the Committee decided to move forward with withdrawing care under the Act, but with the understanding that the Committee and hospital would assist Ms. Hudson with finding and retaining an attorney to represent her in the matter.

The heart-rending nature of the Sun Hudson case suggests that changes to the advance directives and guardianship laws could assist physicians and hospitals in dealing with situations where the caregiver of a person in an end-of-life situation is believed to be mentally incompetent. Because such a scenario can just as easily occur in an adult hospital setting23 as it did at Texas Children’s, potentially all physicians and hospitals could benefit from the changes described below.

As an initial matter, statutes addressing end-of-life care disputes could provide physicians and hospi-

The Committee decided to move forward with withdrawing care under the Act, but with the understanding that the Committee and hospital would assist Ms. Hudson with finding and retaining an attorney to represent her in the matter.
a guardian ad litem who would then be responsible for making decisions on the patient’s behalf, including whether or not to continue with end-of-life treatment. Of course, a problematic consequence of a physician so formally questioning the decisionmaker’s competency would likely be irreparable harm to the physician’s relationship with the decisionmaker; however, ensuring a competent decisionmaker may well be worth this consequence.

In addition, the relevant state family code could give standing to physicians and hospitals caring for patients who are involved in end-of-life disputes to seek court intervention at an earlier stage of the dispute. This would allow physicians and hospitals to voice their concerns even in situations in which CPS, or the equivalent state agency, refuses to get involved. A more radical option under the relevant state family code would be to allow a physician to obtain temporary protective custody of a child, without a court order or consent of the child’s parents, if the physician reasonably believes the child is at risk of abuse or neglect by a child, without a court order or consent of the child’s parents or legal guardians. Similar statutes have been passed in other jurisdictions, and provide physicians with the legal authority to act in the best interests of their patients. In Sun’s case, if such a statute existed, his physicians at Texas Children’s could have taken him into custody after it became clear that Ms. Hudson would not consent to the withdrawal of treatment even when presented with overwhelming evidence that continuing treatment was futile. The physicians’ action would have resulted in CPS and/or the courts becoming involved very early in the process, and could have helped resolve competency issues in a much more streamlined manner than what is now available in states such as Texas. Since the patient’s parents would likely be upset by the physician’s actions, the law should also provide for immunity to physicians who act in good faith. Concerns that physicians may overuse or abuse their power under the statute are reasonable, but statistics from Illinois, where such a statute has been in place for more than twenty years, indicate physicians utilize their power judiciously.

Conclusion

Since the Hudson case ended, some in the legal community have disagreed with the Committee’s decision to assist Ms. Hudson with finding and paying for her legal representation. They argue that Sun’s best interests were paramount in this scenario, and that the delay caused by the Committee’s concerns relating to Ms. Hudson’s competence was unnecessary. With all due respect to the critics’ opinions and analysis on this point, the Committee’s and the hospital’s desire and resolve to ensure that everyone involved in Sun’s case was adequately represented was admirable. It was a case of first impression for the Committee and the hospital, who simply did what they thought was right. Hopefully, other hospitals can learn from our experience and the relevant laws can be amended to help others avoid similar situations in the future.

References

1. The facts set forth in this article are based on the Author’s personal involvement in the case.
4. See Tex. Health & Safety Code Ann. §166.031(2) (“Qualified patient” for withdrawal of care means a patient with a terminal or irreversible condition that has been diagnosed and certified in writing by the attending physician”). A “terminal condition” under the Act is defined as “an incurable condition caused by injury, disease or illness that according to reasonable medical judgment will produce death within six months, even with available life-sustaining treatment provided in accordance with the prevailing standard of care. Tex. Health & Safety Code Ann. §166.002(13); An “ irreversible condition” under the Act is defined as “a condition, injury, or illness (A) that may be treated but is never cured or eliminated; (B) that leaves a person unable to care for or make decisions for the person’s own self; and (C) that, without life-sustaining treatment provided in accordance with the prevailing standard of medical care, is fatal.” Tex. Health & Safety Code Ann. §166.002(9).
5. See MedicineNet.com, Definition of Dysplasia, Thanatophoric, at <www.medterms.com/script/main/art.asp?article key=9729> (last visited September 14, 2005) (“Thanatophoric dysplasia, also called thanatophoric dwarfism, was discovered in 1967 by Pierre Maroteaux and his coworkers who used the Greek term “thanatophoric” meaning death-bringing.”). See K. Baker, D. Olson, C. Harding, and R. Pauli, “Long-Term Survival in Typical Thanatophoric Dysplasia Type 1,” American Journal of Medical Genetics 70 (1997): 427-36 (stating most affected individuals die of respiratory insufficiency in the first hours or days of life, with the respiratory insufficiency being linked to a small chest cavity and lung hypoplasia).
16. See Tex. Health & Safety Code Ann. §166.035. (“The following persons may execute a directive on behalf of a qualified patient who is younger than 18 years of age: (1) the patient’s spouse, if the spouse is an adult; (2) the patient’s parent; or (3) the patient’s legal guardian.”). One would logically assume that only the decisions of competent decisionmakers carry weight under the Advance Directives Act, but nowhere within the statute is that assumption expressed.
21. Wanda Hudson, as the Mother of Sun v.
Texas Children's Hospital, 2004-66483 (333rd Judicial District, Harris County, Texas, filed November 23, 2004). For jurisdictional reasons, Ms. Hudson’s attorney subsequently dismissed this case filed in the State District Court and re-filed in the Harris County, Texas, Probate Court. See Wanda Hudson, as the Mother of Sun v. Texas Children’s Hospital, No. 352,526 (Probate Ct. No. Four, Harris County, Texas, filed Dec. 29, 2004).

24. Tex. Health & Safety Code Ann. §166.046(g); Final Order Denying Plaintiff’s Request for an Extension of Time Under the Health & Safety Code Section 166.046(g) and Denying Plaintiff’s Motion for a New Trial and to Reconsider Ruling Granting Defendant’s Special Exceptions. Wanda Hudson, as the Mother of Sun v. Texas Children’s Hospital, No. 352,526 (Probate Court No. Four, Harris County, Texas, March 14, 2005).

25. It is quite easy to imagine a scenario where an elderly man has a stroke and afterwards is deemed to have an irreversible condition under the Texas Advance Directives Act but where his wife, also elderly, is suffering from early dementia and gets confused whenever anyone speaks to her about her husband’s medical condition, particularly if the couple has no children or other relatives available to assist in the decision-making process.


27. See, e.g., 325 Ill. Comp. Stat. 5/5 (“An officer of a local law enforcement agency, designated employee of the Department, or a physician treating a child may take or retain temporary protective custody of the child without the consent of the person responsible for the child’s welfare, if (1) he has reason to believe that the child cannot be cared for at home or in the custody of the person responsible for the child’s welfare without endangering the child’s health or safety; and (2) there is not time to apply for a court order under the Juvenile Court Act of 1987 for temporary custody of the child”); Iowa Code § 232.79 (allowing a physician to take the child into custody if the child is in a circumstance or condition that presents an imminent danger to the child’s life or health, and, if there is not enough time to apply for an order).

28. See, e.g., 325 Ill. Comp. Stat. 5/5 (providing that “[a]ny physician authorized and acting in good faith and in accordance with acceptable medical practice in the treatment of a child under this Section shall have immunity from any liability, civil or criminal, that might otherwise be incurred or imposed as a result of granting permission for emergency treatment”).

The Institute of Medicine’s 2005 publication, *Dietary Supplements: A Framework for Evaluating Safety*, is authoritative and thorough, and thus representative of other reports by the Institute of Medicine (IOM). What makes this report particularly interesting, however, is the rich political subtext that exists in the interstices of the report, popping up here and there in brief comments and barely suppressed yelps of exasperation. To understand this context, it is useful to reflect for a moment on the special nature of the IOM and its relationship to government.

IOM is part of the National Academy of Sciences, and is a private, non-governmental organization that does not receive direct federal funding for its work. Rather, IOM studies are often funded by contract with governmental entities that request reports on particular topics. As a case in point, the dietary supplement study was requested and paid for by the FDA.

IOM is quite particular about the process it uses to generate a report. IOM gathers a group of subject matter experts who serve as the report’s authors. The requestor of the report may suggest experts or types of expertise, but IOM retains control of the selection of the committee, and these experts serve voluntarily. It is a tribute to the respect with which IOM is regarded that highly skilled professionals are willing to serve on its committees for free, or rather only for the compensation provided by the associated prestige. Committees reviewing a particular topic attempt to reach consensus, relying upon scientific evidence wherever possible, and noting where they must rely upon expert opinion in the absence of conclusive evidence. Once the initial draft is complete, it is submitted to expert reviewers whose comments are anonymous to the committee until after revisions are complete. IOM staff follow a formal review process, in which they track all reviewer comments and incorporate the comments into the report, or provide a rationale for rejecting the reviewer’s advice. IOM states on its website that sponsors do not have the right to review a report before publication or to edit the conclusions.

Why bother to review IOM’s process in such detail? Because this approach is specifically designed to insulate scientific expertise from manipulation by the political process and special interest groups. The firewall that exists between expertise and politics is the bedrock of IOM’s prestige. The process may not be perfect, but it remains the gold standard against which other governmental reports are measured.

What then is the nature of the firewall that the IOM is asked to construct in its report on dietary supple-
ments? The report is commissioned by the FDA, which requested a framework for evaluating the safety of dietary supplements. But surely the FDA knows how to evaluate safety! FDA's 10,000 employees evaluate the safety of many sorts of products, from drugs to medical devices to cosmetics. Why should the task of evaluating safety be any different for dietary supplements than it is for other products that FDA evaluates? In fact, FDA's task in evaluating dietary supplements is exceedingly difficult, not only because evaluating safety is always complex, but also because the regulatory straightjacket that dictates FDA's actions regarding dietary supplements poses significant obstacles. FDA's request to IOM amounts to the question: how can we evaluate safety given these political constraints? IOM's response offers a sober review of appropriate methods given the current circumstances. It also notes that these circumstances demand significant legislative revision.

The political context of this report has everything to do with DSHEA, the Dietary Supplement Health and Education Act of 1994. This interesting piece of legislation enjoyed strong bipartisan support, particularly from Senators Hatch and Harkin. The act is the principle legislative revision. Its wisdom of giving FDA the authority to impose warning labels on drugs when appropriate. For dietary supplements, FDA has no such authority.

In sum, FDA cannot demand safety information for dietary supplements in advance of sales, cannot demand post-market information from makers about product-related harm, and cannot impose warning labels on dietary supplement products when it learns about significant risks. Dietary Supplements documents that FDA's toolkit is thus limited for evaluating the safety of dietary supplements, and strives for a neutral tone. However, phrases referring to the "unique regulatory context" cannot diminish the overwhelming impression that this context is not so much unique as unworkable.

The question then that FDA puts before IOM is how to evaluate dietary supplements within the stated constraints. IOM's proposal is thorough and rational; it reviews various types of information that are likely to be available for different dietary supplements and articulates a plan for weighing this evidence in order to assess safety. Separate chapters analyze human data, animal data, in vitro data and data from related substances. Along the way, the reader gleams numerous useful facts about the general scientific approach to evaluating safety. For instance, chapter 3 offers a useful summary of science-based principles, and directly addresses the argument, often supplied by advocates of dietary supplements and complementary medicine, that historical use of a...
The report is rich with specific examples, some of which produce an alarming sense of the high level of data that are required before FDA may take action against a dietary supplement. The report notes that “historical use should not be used as *prima facie* evidence that the ingredient does not cause harm.”

The report is rich with specific examples, some of which produce an alarming sense of the high level of data that are required before FDA may take action against a dietary supplement. Chapter 8 discusses St. John's wort in a passage on interactions between dietary supplements and medications, noting the oft-cited capacity of this supplement to decrease effective drug levels, including for life-saving treatments like cyclosporine, which must be maintained at or above a certain level to prevent rejection of transplanted organs.

More disturbingly still, the prototype monograph on chaparral summarizes available information on possible toxic effects on the liver, kidney and reproductive organs. The lack of data regarding dietary supplements is frequently cited as a barrier to more effective regulation. However, this report makes clear that safety data for various supplements already exist that would likely lead to restrictive actions if they applied to a drug, food additive or other class of product. We do not entirely lack data on dietary supplements, but the data we have do not suffice to overcome strictures imposed by DSHEA.

Nor does *Dietary Supplements* rely on the most egregious case of a hazardous supplement that FDA attempted to restrict. Perhaps believing that by publication time a federal ban would conclusively remove ephedra from the shelves of American stores, this report barely mentions ephedra in its 400 plus pages; ephedra does not appear in the book's index. The omission is unfortunate, for the example of ephedra is perfectly designed to reveal the frustrating impact of DSHEA on the FDA's attempts to regulate dietary supplements. More than ten years of effort, spurred by convincing reports of deaths associated with ephedra products, culminated with FDA's published ruling banning the sale of ephedra products. Recently, a federal court undermined that ruling by upholding a challenge to it, perhaps re-opening the door for the legal sale of low-dose ephedra products.

The concluding chapter recommends various changes to DSHEA, many of which have been proposed in previous governmental reports. Recommendations include submission of all pre-market data that manufacturers possess, mandatory adverse event reports from manufacturers, and new authority to impose warning labels on dietary supplements when deemed appropriate. The authors stepped beyond the bounds of FDA's explicit request in offering these recommendations, and are to be commended for doing so. Adopting these recommendations would greatly promote the goal of the efficient and thorough evaluation of dietary supplement safety.

*Dietary Supplements* lays out a detailed description of how to write a monograph; one modest criticism of the report is that this section in particular and the use of language throughout the report produce a slightly odd and doubtless unintended effect. On the one hand, some aspects of the process are explained in the minutest detail, down to the provision of a definition for the word “monograph.” Why, the reader wonders, should someone at FDA who does not know what a monograph is be employed in writing one? On the other hand, some fairly arcane technical knowledge is assumed; 50-cent words like “pharmacognosy” and “pharmacoepidemiology” go undefined, leading the reader to wonder about the intended audience for this report. Perhaps the authors are replying to FDA colleagues who already know the answers, but need them presented by experts to quiet those who doubt the need for a vigorous evaluation of dietary supplement safety. Or perhaps the report is to be taken more literally, and simply wishes to answer in explicit detail the question of how one should evaluate the safety of dietary supplements. In any case the result is oddly similar at times to the passage in *One Hundred Years of Solitude* in which villagers attempt to write down the method for doing familiar tasks, such as milking a cow, before they are all overcome by amnesia. The reader wonders if it is really necessary to tell the FDA how to weigh evidence and write a monograph; the exercise is not likely to engage the general educated reader. This criticism is intended as a minor one, however. The report's explicit purpose is to articulate technical procedures for professionals; like a sonnet it adheres to its designated format. *Dietary Supplements* is not intended for citizens, or at least not for those who do not write monographs, but for their public servants who do, and it fulfills that mission admirably. Though this volume will not knock *The Sisterhood of the Traveling Pants* from the best-seller list, it will and should be read by those at FDA charged with the challenging task of evaluating supplements in the context of DSHEA. The report's thoroughness may lend it an appeal for professionals beyond FDA as well. State-level entities and academics working on dietary supplements will find a useful template for scientific enquiry in a field that is often riddled with sub-standard research. *Dietary Supplements* is a commendable response to a question formed in the crucible of science and politics.

Barbara A. Noah

In the past fifteen years, dietary supplements, including vitamins, minerals, herbs, and amino acids have become increasingly popular with consumers, resulting in a $20 billion industry.1 These products create significant regulatory challenges for the Food and Drug Administration, (“FDA”) the federal agency charged with supervising the safety of these products. In 1993, the FDA published a notice that summarized safety concerns associated with various categories of dietary supplements based on the recommendations of an agency task force, and proposed some rather aggressive regulatory strategies.2 In 1994, apparently in response to the 1993 notice, Congress enacted the Dietary Supplement Health and Education Act (“DSHEA”),3 which provides the FDA with limited authority to regulate covered products.

In drafting DSHEA, Congress chose not to create an entirely new category of products subject to FDA controls. Instead, Congress accomplished its deregulatory goals by declaring that dietary supplements constitute a subcategory of food.4 DSHEA explicitly exempts these products from regulation as food additives, an important concession because food additives are subject to pre-market approval requirements.5 Congress also opted against treating dietary supplements as drugs under the Food, Drug, & Cosmetic Act (FDCA).6 In contrast to the regulatory scheme for drugs, which requires substantial pre-market evaluation of safety and efficacy before the granting of a license,7 DSHEA allows dietary supplement manufacturers to market their products without receiving any pre-market clearance from the FDA.8

In June 2005, just over a decade after the passage of DSHEA, the New York State Task Force on Life & the Law published a comprehensive report dealing with dietary supplement safety entitled Dietary Supplements: Balancing Consumer Choice & Safety (the “Task Force Report” or “Report”).9 As the Preface to the Report explains, the Task Force was created by the state’s governor in 1985 and serves as a policy development and educational resource for the New York State government, focusing primarily on issues at the intersection of health and law. The Task Force Report makes a strong case for the following propositions: that dietary supplement use poses significant risks, that existing FDA authority over supplements is inadequate to guard against these risks, and that the public is ill-equipped to make informed choices about the use of these products because of limited available information about product risks and benefits. Based on these observations, the Task Force Report recommends more aggressive state regulation of supplement products.

The Task Force Report begins with a case study in dietary supplement safety. In Chapter One, the Task Force describes and evaluates the safety problems associated with dietary supplements containing ephedrine alkaloids. As the Report explains, DSHEA does not require pre-market safety or efficacy evaluation of dietary supplements. Instead, the statute requires the FDA to prove evidence of a supplement’s risks in order to declare it adulterated and order its removal from the market.10 Between 1996 and 2003, several individuals died after ingesting ephedra-containing products. For example, in 1999, 21-year-old John Leseman died during exercise while taking an ephedra product and, in 2003, 23-year-old Baltimore Orioles pitcher Steve Bechler died after collapsing during spring training; he too had been taking an ephedra-based weight loss product.11 Problems with the ingredient surfaced much earlier. Beginning in 1994, the FDA issued a series of medical bulletins and consumer alerts warning of the risk of adverse effects from ephedra supplements.12 By 1997, the agency had received over 800 adverse event reports concerning ephedra, and some estimates suggest that this ingredient played a role in at least 155 deaths.13

Based on these developments, the FDA concluded that such supplements met DSHEA’s adulteration standard, presenting a significant or unreasonable risk of illness or injury.14 Initially, the FDA considered addressing these safety concerns with restrictions on dosage and recommended duration of use,15 but the agency ultimately concluded that only a complete ban on ephedra supplements would adequately protect consumers.16 Seven years after the initial safety concerns arose, in 2004, the FDA issued a final rule declaring ephedra products adulterated and ordered all sales of products containing the ingredient to cease.17 As the Task Force Report explains, a year later, a federal district court in Utah struck down the FDA rule as applied to ephedrine products containing 10 milligrams or fewer of active ingredient.18 The court opined that the FDA’s interpretation of DSHEA’s “unreasonable risk” standard, which involved a risk-benefit analysis, was inconsistent with Congressional intent because it “places a burden on [manufacturers] to demonstrate a benefit as a precondition of sale, and that is contrary to Congress’ intent.”19 It is not clear at this point how the court ruling will impact products that contain more than 10 milligrams of active ingredient.20 The Task Force Report appropriately devotes a good deal of space to the ephedra case study because it effectively illustrates the deficiencies in the current federal regulatory system. As the Report explains, the FDA is ill-equipped to regulate dietary supplements because it lacks authority to demand pre-market proof of safety, or to require post-market adverse event reports.21 For these reasons, the Report endorses state regulatory action to fill federal regulatory gaps. New York took precisely...
this approach in October 2003, when it banned supplements that contain ephedrine alkaloids.

In Chapter Two, the Task Force examines the prevalence of and reasons for consumer dietary supplement use. After reviewing inconsistent data from various polls and studies about dietary supplement consumption, the Task Force Report observes that it is difficult to estimate utilization rates for different types of supplements, probably because many consumers purchase supplements via the Internet, making marketing data inaccurate.22 Existing data suggest, however, that dietary supplement consumption varies by gender and race and that supplement use is generally on the increase.23 This chapter also includes a discussion of common rationales for supplement use for reasons that extend well beyond mere supplementation of the diet. Typical consumers of supplements take these products to ensure good health, enhance physical appearance or performance, and to treat both benign medical problems like the common cold and serious medical conditions such as arthritis, cancer, and prostate enlargement.24 Interestingly, the data suggest that many consumers view dietary supplements as safe and effective substitutes for prescription drugs despite the lack of clinical evidence to support such uses,25 a point that the Task Force relies on later in support of the its recommend state regulatory reforms.

In Chapter Three, the Report examines the evidence of potential benefits and risks associated with a variety of dietary supplements.26 Because the federal statute does not require premarket evaluation of safety or efficacy for dietary supplements, that same mindset would affect randomized, controlled trials of dietary supplements. The clinical research data on potential herb-drug interactions suffer from the same flaws. Because DSHEA does not require premarket proof of safety or efficacy, researchers have little interest in conducting such trials, and funding for this work remains limited.29 Instead, manufacturers and consumers of supplements mistakenly believe that a history of use for a particular purpose serves as a proxy for clinical evidence of safety and efficacy.30 Because they are pharmacologically active products, dietary supplements both offer potential benefits and create genuine risks. The Chapter reviews evidence of documented benefits associated with certain supplements, such as folic acid to prevent neural tube defects and the ingestion of vitamin supplements to ensure adequate intake of essential nutrients.31 As the Report observes, lack of evidence of harm does not necessarily indicate safety, and even those products which are apparently both safe and efficacious may harm consumers who choose to self-diagnose and treat, thereby delaying conventional medical care.32 The Task Force Report could be more precise on these points. It tends to collapse its discussion of safety and efficacy into a single argument without explicitly recognizing that this implies that risk-benefit balancing for dietary supplements is appropriate. In fact, because supplements are regulated as a category of food, and because the traditional food safety standard does not countenance risk-benefit balancing,33 it is far from obvious that risk-benefit balancing makes sense for supplements. Nevertheless, the FDA appears to have embraced risk-benefit balancing for these products, though its tolerance for risk appears quite low.34 Moreover, like other pharmacologically active products such as prescription drugs, questions of safety and efficacy require independent evidence: a particular supplement may be safe to ingest but not efficacious, or it may be efficacious for a particular use, but not safe. The deficiencies of clinical evidence discussed in the chapter become even more pronounced when one considers the complexity of the evidence required to make a truly meaningful assessment of the utility of a dietary supplement product.

In addition, the Report explains that dietary supplements may interact negatively with a variety of traditional pharmaceutical products, interfering with drug efficacy or resulting in changes in the way the body metabolizes various drugs.

The Report explains that dietary supplements may interact negatively with a variety of traditional pharmaceutical products, interfering with drug efficacy or resulting in changes in the way the body metabolizes various drugs.
Chapter Four focuses on the federal regulation of dietary supplements. The Chapter begins with an early history of FDA regulation which, although not essential to an understanding of the current issues, provides some context for the discussion of federal dietary supplement regulation. In its discussion of DSHEA, the Task Force provides additional detail about the differences between a structure or function claim and a disease claim. DSHEA permits manufacturers to include so-called structure or function claims in their product labeling, as long as the manufacturer “has substantiation that such statement is truthful and not misleading.” and it does not require that manufacturers pre-clear these claims with the FDA. By contrast, such claims on a drug product’s label would render the product an unapproved new drug in the absence of explicit FDA pre-approval. If dietary supplement manufacturers wish to include structure or function claims on their product labels, they need only add a disclaimer that the product has not been evaluated by the FDA and that the “product is not intended to diagnose, treat, cure, or prevent any disease.” Consumers may fail to appreciate the distinction between a true therapeutic claim and a structure or function claim, however, and may misinterpret or ignore the disclaimer.

There is an important problem with the current dietary supplement regulatory scheme that is outlined in chapter four. Despite the fact that DSHEA generally does not permit disease claims and limits supplement labels to so-called “structure or function” claims, the Task Force Report explains that many consumers use dietary supplements not simply to supplement their diets but also for drug-like purposes: to ensure good health, enhance physical appearance or performance, and to treat both benign and serious medical problems. As the Report points out, consumers mistakenly believe that the nonprescription status of supplements and their “naturalness” make these products inherently safe. Moreover, consumers incorrectly assume that a government agency pre-approves supplements, permits claims about safety and effectiveness only with the support of sound scientific evidence, and requires warnings about potential side effects. Finally, the Task Force Report observes that, because physicians and most dietitians receive limited training about the proper use of supplements, and because internet information can be incomplete or deceptive, many consumers remain ill-informed about the risks of these products.

The remainder of the chapter focuses on two key issues – adverse event reporting and Federal Trade Commission enforcement of advertising and marketing standards. As explained above, DSHEA does not authorize the FDA to require adverse event reports relating to dietary supplements, and estimates suggest that the FDA receives reports about less than one percent of all adverse events associated with supplement products. Because the statute fails to mandate reporting, some manufacturers simply choose not to collect data on adverse events, while others collect it but fail to report it to the FDA. Noting that poison control centers and health care providers also could cooperate to improve voluntary adverse event reporting, the Task Force Report nevertheless urges mandatory dietary supplement reporting at the federal level. The remainder of the chapter explores the division of labor between the FDA, which retains primary responsibility for supplement labeling, and the FTC, which enforces advertising and marketing regulations, and describes several recent cooperative enforcement actions against individual supplement products and manufacturers.

In chapter five, the Task Force briefly considers some of the legal issues raised by concurrent state regulation of dietary supplements, along with private sector initiatives to improve dietary supplement safety. The chapter begins with a brief summary of federal preemption doctrine, concluding that DSHEA and a related federal nutrition labeling law leave New York State “with significant retained authority to regulate supplement labeling and marketing.” This section of the Task Force Report also provides an overview of the various departments within the New York State government that might play a role in more aggressive state regulation of dietary supplements and describes several past state retail restrictions and product seizures, including bans on the sale of twenty-six different herbal products containing ephedrine alkaloids.

The Task Force Report concludes in chapter six with a series of recommendations to improve the quality of information about dietary supplement safety and efficacy for consumers of these products and to empower the New York Department of Health to respond quickly and efficiently when scientific evidence suggests that a particular supplement product or ingredient poses genuine risks to consumers. The Task Force acknowledges the common concern that individual state regulatory efforts will create a patchwork of regulations that the industry may find burdensome, but concludes that, until the federal regulatory scheme permits more effective FDA responses, state intervention is necessary to protect the public.

The Task Force recommends that the State Commissioner of Health take two major steps: (1) create an Expert Committee within the Department of Health to collect data and evaluate dietary supplement safety and to make recommendations based on the data and; (2) undertake a major public health education campaign about the risks and benefits of dietary supplements. As part of the first recommendation, the Task Force envisions that the Expert Committee will serve as a central repository of safety information about dietary supplements and will facilitate federal regulatory responses by sharing all available data with the FDA. The Task Force also recommends that the Expert Committee institute mandatory adverse event reporting by dietary supplement manufacturers and distributors doing business in New York
State, at least until such time as Congress amends DSHEA to require such reporting nationally.55

In order to enforce the mandatory reporting requirement, the Task Force Report also suggests that the Expert Committee create a registry of New York manufacturers and distributors, and require a registration fee to fund administrative costs.56 Along with its collection and processing of safety data, the Expert Committee could require labeling changes to reflect new information or even recommend that the sale of certain products be banned, though, as the Task Force Report acknowledges, such a measure would require explicit authority from the State Legislature or the Commissioner of Health.57 Finally, the Task Force urges the State Department of Health to undertake a public education campaign in order to inform consumers about the risks and benefits of supplement use, the risks of supplements when used in combination with prescription drugs, and the process for and importance of reporting adverse events associated with these products.58

Although the general description of federal preemption doctrine in chapter five appears accurate, some of the proposed state regulatory initiatives nevertheless could encounter preemption difficulties. Despite the fact that DSHEA does not contain an express preemption clause, it is possible that some of the more pro-active recommendations in chapter six (particularly the proposals to require manufacturer reporting of serious adverse events and to require manufacturers to register and to pay a fee) might pose “conflict in purpose” implied preemption problems, since the “purpose” of DSHEA is clearly deregulatory. For example, conflict of purpose preemption would, at least in theory, prevent the State of New York from banning structure or function claims entirely on the labeling of dietary supplements (because DSHEA expressly permits such claims).59 In contrast, the Task Force Report’s proposal to mandate the submission of safety information in the form of adverse event reports appears less obviously in conflict with the deregulatory spirit of DSHEA (compared with, for example, a proposal to mandate pre-market safety data), although one could argue that even mandated adverse event reporting conflicts with the “spirit” of DSHEA. Of course, if the FDA were to implement a federal adverse event reporting requirement for dietary supplements down the road, it appears possible that it would likely preempt a state reporting requirement, but for now that is a hypothetical problem.

The Task Force Report’s proposal to require a safety/efficacy disclaimer overlapped with the federal disclaimer requirement; because there is a directly on-point federal requirement, the proposed state requirement appears somewhat duplicative. Having said that, because DSHEA does not contain an express preemption clause, it does not seem that the state disclaimer requirement would run into problems under implied preemption principles because the proposed state disclaimer is not in tension with the federal requirement. With respect to proposed state power to require warnings about particular products, since DSHEA is silent about warning labels, there is probably little risk of federal preemption.60 This overview of possible preemption issues is necessarily speculative, but members of the Task Force and the State legislature might consider exploring the preemption questions further prior to implementing the Report’s proposals.

Overall, the Task Force Report paints a persuasively grim picture about dietary supplement safety regulation. Although the Task Force steers clear of one of the most complex questions underlying the DSHEA safety scheme – that is, whether risk-benefit balancing truly is appropriate for products which are, nominally at least, a type of food, and for whom sound evidence of safety and efficiency are lacking – the Task Force Report makes a convincing case for the proposition that consumers need better information in order to make good choices about supplement use. Implicitly, the Report suggests that, with adequate information about the safety and efficacy of these products, the process of risk-benefit balancing ought to be left to the individual consumer. This position presupposes that consumers will routinely comprehend and consider such information if it should be made available to them. Given what the literature suggests about the effectiveness of informed consent to ensure quality decision-making in other health care settings,61 such a proposition appears unduly optimistic.

References
2. FDA, Advance Notice of Proposed Rulemaking on Dietary Supplements, 58 Fed. Reg. 33,690 (1993) (recommending that the FDA establish safe levels for use of vitamins and minerals, that it regulate drugs amino acid-containing dietary supplements as drugs, and that various other types of supplements be evaluated under the food provisions of the statute in order to determine whether they are generally recognized as safe [GRAS]).
4. “Definition of certain foods as dietary supplements, 21 U.S.C. § 321, is amended by adding at the end the following: The term dietary supplement: (1) means a product (other than tobacco) intended to supplement the diet that bears or contains one or more of the following dietary ingredients: (A) a vitamin; (B) a mineral; (C) an herb or other botanical; (D) an amino acid; (E) a dietary substance for use by man to supplement the diet by increasing the total dietary intake; or (F) a concentrate, metabolite, constituent, extract, or combination of any ingredient described in clause (A), (B), (C), (D), or (E).” 21 U.S.C. § 321(g). The statute further defines dietary supplements as products that are labeled as “dietary supplements” and are not represented for use as a conventional food or as a sole item of a meal or the diet. See 21 U.S.C. § 321(ff).
5. 21 U.S.C. § 321(g). The FDA subjects food additives to pre-market review in order to determine whether such substances are safe for use. See id. § 348(a)-(c). The FDA has, however, exempted from pre-market approval requirements substances added to food that are generally recognized as safe (“GRAS”). See 21 C.F.R. § 170.30(a) (2004); also see generally L. Noah and R. A. Merrill, “Starting from Scratch? Reinventing the Food Additive Approval
3. 21 U.S.C. § 321(g)(C). The definition provisions specifically exempt foods and dietary supplements from regulation as drugs, even if their labeling contains certain types of health claims permitted under DSHEA. See id. § 321(g)(1)(D).
5. A company wishing to sell a supplement containing a “new dietary ingredient,” (defined as one not marketed before October 15, 1994) however, must file a notification with the FDA at least 75 days prior to market introduction, which provides the basis for the manufacturer’s conclusion that the supplement will “reasonably be expected to be safe” and must demonstrate only that “[t]here is a history of use or other evidence of safety.” 21 U.S.C. § 350b(2). If the agency finds the notification inadequate, it can prevent marketing only by initiating formal enforcement proceedings, under the provisions of DSHEA.
9. Ephedra can cause a variety of serious adverse events, including irregular heartbeat, sleeplessness, anxiety, tremors, headache, seizures, heart attack, and stroke. See Task Force Report, supra note 9, at 16.
10. Task Force Report, supra note 9, at 16-17; see also T. Hampton, “More Scrutiny for Dietary Supplements,” *JAMA* 293 (2005): 27-28, at 28 (noting that critics argue that the FDA waited too long to act, promulgating the final rule declaring ephedra products adulterated only after reports of at least 155 deaths associated with the product’s use).
12. FDA, Proposed Rule, 62 Fed. Reg. 30,678 (1997) (proposing new rules that would render dietary supplements containing ephedra adulterated if they contain more than 8 mg. of ephedrine alka-
regulatory authority to require substantiating data from manufacturers. See id. at 56.  
41. 21 U.S.C. § 321(g)(C) (“The term ‘drug’ means…articles (other than food) intended to affect the structure or function of the body of man…”); see also FDA, Structure/Function Claims Small Entity Compliance Guide, available at <http://www.cfsan.fda.gov/~dms/sclmguid.html> (last visited September 15, 2005) (providing examples of permissible claims and advice about divining the difference between structure and function claims and impermissible disease claims).  
42. 21 U.S.C. § 343(r)(6)(C) (2000). As the Task Force Report elaborates in Chapter Four, the claim that “Calcium builds strong bones” is an example of a structure/function claim that would not require FDA approval on a dietary supplement label, whereas a claim that “Calcium prevents or reduces the effects of osteoporosis” would qualify as a specific disease claim requiring FDA approval. See Task Force Report, supra note 9, at 56.  
43. Task Force Report, supra note 9, at 27-30.  
44. Id. at 30. On a related note, chapter four also explains that DSHEA grants the FDA’s authority to issue regulations concerning good manufacturing practices (“GMPs”) for supplement products, noting that the current GMPS, which have yet to become final, are modeled on food GMPs rather than those for pharmaceutical products. As the Task Force Report explains, the proposed GMPS fail to address two key issues-potential efficacy and toxicity of ingredients, focusing instead on reducing the risk of harm due to super-potency or contaminants. See id. at 40.  
45. Id.  
46. Id. at 60 (explaining that the FDA received only 10 adverse event reports from supplement manufacturers between 1994 and 1999 despite the fact that approximately 100 million U.S. consumers used these products).  
47. Id. at 60-61 (describing information gaps in those adverse event reports that the FDA receives).  
48. Id. at 63.  
50. Task Force Report, supra note 9, at 70.  
51. Id. at 71-72.  
52. Task Force Report, supra note 9, at 78.  
53. Id. at 79-80.  
54. Id. at 79-80 (recommending that the Expert Committee use federal regulatory standards for defining “serious adverse events” and urging voluntary reporting by consumers and health care professionals of less serious adverse events in order to facilitate identification of longer-latency toxicity problems).  
55. Id. at 80-81.  
56. Id. at 81.  
57. Id. at 83.  
58. The Task Force does not explicitly propose such a change in labeling policy but does discuss the prospect of mandated labeling of specific risks for specific products and blanket labeling of risks to pregnant and lactating women. See Task Force Report, supra note 9, at 81.  
59. Committee of Dental Amalgam Mfrs. v. Stratton, 92 F.3d 807 (9th Cir. 1996) (holding that California’s Prop. 65 safety warning for carcinogenic products were not preempted by the federal statute).  

CROWN LIFE BIOETHICS FELLOW 2006-2007  

The Institute for Bioethics, Health Policy, and Law at the University of Louisville School of Medicine is seeking applications for a bioethics fellow for the academic year 2006-2007. Funded by a grant from the Crown Life Insurance Company of Canada, the fellowship is available for a recent graduate with a terminal academic or professional degree who is interested in pursuing a career in bioethics research.  

Applicants should have a Ph.D., J.D., M.D., Dr. P.H. or comparable degree. Although the position is not limited to a specific subject matter, the Institute has recently concentrated on research ethics, public health ethics, and genetics. The fellow will engage in a range of research activities, both individually and in collaboration with Institute faculty. Teaching and lecturing opportunities will be available, and the fellowship comes with a travel stipend to support professional development. Compensation will be based on the NIH scale for post-doctoral fellows.  

Nominations and applications should include a CV and a cover letter setting forth research interests and possible plans for the fellowship year, and how the fellowship will advance their career. They should be addressed to: Professor Mark A. Rothstein, Director, Institute for Bioethics, Health Policy, and Law, 501 East Broadway Suite 310, Louisville, KY 40202.
For years, legal scholars and environmental activists have maintained that traditional tort proof requirements create insurmountable obstacles to recovery for most plaintiffs in chemical exposure cases, be they pharmaceutical suits or environmental toxic tort cases. Generally, tort law requires a plaintiff to show that the defendant owed a duty, that the defendant breached that duty, and that the breach of that duty caused the injury that is the subject of the suit. In some cases those requirements can be relaxed, as for example, when an injury is of a type that does not occur except as a result of negligence (under the doctrine of res ipsa loquitur). In most toxic tort and pharmaceutical cases, the central issue is proof of causation, because the standard for admissibility of scientific evidence can be very high and such evidence is expensive and often unavailable.

In the recent Vioxx case, however, the lack of evidence of causation was not an obstacle to a very large damage award. The case turned on juror impressions of Merck’s corporate behavior. Jurors reportedly believed that Merck had acted unscrupulously in marketing a drug it knew to be dangerous, and that belief overcame the causation requirement. On appeal, Merck will argue that the causation requirement was not met, but in its other Vioxx cases, it seems poised to shift its tactics. Reports of current trials suggest that instead of trying to refute evidence of causation, Merck is attempting to refute evidence that it knew the drug to be dangerous. In doing so, Merck is acknowledging that the battleground appears to be the morality of the company’s actions, not the validity of the injury claim.

Factual and Legal Background
Robert Ernst, a marathon runner, fitness trainer, and assistant produce manager for Wal-Mart, died of a cardiac arrhythmia and atherosclerosis after taking Vioxx for seven months. In her suit against Merck filed in Texas’s 23rd District Court in 2002, his wife, Carol Ernst, alleged negligent failure to warn of the dangers associated with the drug and civil conspiracy to conceal the danger. By the time the case went to trial, it was one of about 4,200 Vioxx suits pending nationwide.

The Ernst case seemed favorable for Merck because of state-law limitations on liability and the facts of the case. First, Texas law caps punitive damages at twice the amount rewarded for economic damages. Second, non-economic damages were limited because Carol and Robert Ernst had been married for less than one year. Economic damages were likely to be small because Ernst’s income was small and he did not suffer a prolonged illness. According to Merck’s lawyers, however, the most favorable factor was the lack of evidence of causation.

Kent Jarrell, spokesman for Hughes, Hubbard & Reed, one of the law firms representing Merck, said before the trial, “This is not a normal Vioxx case, which is about heart attacks and strokes being increased by the use of the drug...[The plaintiff] is going to make this about the behavior of the company. We plan on making it about the specifics of causation.” Those specifics, as described in Merck’s Background Science Brief, are that “no reliable scientific evidence has ever demonstrated that Vioxx causes cardiac arrhythmias.”
and that no epidemiological study with a placebo had demonstrated a statistically significant effect from Vioxx taken for less than 18 months. The information in that brief supported a motion for summary judgment which read, “Plaintiffs must establish to a reasonable degree of medical certainty that Vioxx probably – not possibly caused Mr. Ernst’s fatal cardiac arrhythmia. They cannot do so. The failure of proof is fatal to each and every claim.”

Before the judge ruled on the motion for summary judgment, however, the legal climate changed. Just 11 days before the Ernst trial was scheduled to begin, Texas Attorney General Greg Abbott filed charges against Merck for falsely advertising the safety of Vioxx and defrauding the Texas Medicaid program. That suit attracted considerable public attention, leading Merck to file for a continuance, arguing that the publicity would prejudice any possible jury and prevent a fair trial. The case did go to trial, though, and played out in an untraditional way.

The Law: The Causation Requirement
For a plaintiff alleging injury by exposure to a toxic chemical (including a drug), the burden of proof requires that he show that he was exposed to the chemical, that the chemical is capable of causing his injury (general causation), that the exposure caused his injury (specific causation), and that the particular defendant was responsible for the exposure that caused the injury. Proving general causation involves introducing statistical evidence comparing the incidence of an injury in the general population and the incidence of the injury in the population exposed to the chemical. Proving specific causation involves testimony that the substance in question was produced by the defendant and probably caused the injury. Both of these proof requirements involve scientific and statistical evidence, the admissibility of which can be controversial.

Since the 1993 decision in William Daubert et al. v. Merrell Dow Pharmaceuticals, Inc, the standard for admitting scientific evidence in federal court has been a two part test: whether testimony conveys scientific knowledge and whether it is relevant. Justice Blackmun explained the first part of the test, saying, “The adjective ‘scientific’ implies a grounding in the methods and procedures of science. Similarly, the word ‘knowledge’ connotes more than subjective belief or unsupported speculation.”

As plaintiffs’ lawyer W. Mark Lanier said before the trial, “I win the case whether it is a heart attack or arrhythmia. Merck internally knew that Vioxx caused both. Merck knew it before it sold its first pill. I don’t have to use outside experts. I will use Merck’s own documents, their own emails, their own scientists.”
far as to call the requirement “antithetical to tort law objectives.”26 The goals of the tort system, to provide individual justice and to deter dangerous activity, may be better served without requiring plaintiffs to prove causation.27

The Law as Applied by the Jury in the Vioxx Case

It appears, from reports and legal analysis, that the jury in Ernst v. Merck & Co. agreed that proof of causation should not be required. There was no question that Mr. Ernst had died, but did Vioxx cause his death? During the Ernst trial, Steven Goode, professor of evidence at the University of Texas School of Law, described Merck’s probable tactics in the context of traditional Daubert debates. He said, “The name of the game for the defense is to knock out the expert witnesses on causation and then move for a summary judgment or directed verdict...If the expert is making too great a leap from the data to his or her conclusion, then the testimony might be viewed as scientifically unreliable and inadmissible.”28 That tactic is a defense though, and depends on the plaintiff putting on a causation case in the first place.

Plaintiffs made it clear from the beginning that they would not make such a case on causation. Instead, as plaintiffs’ lawyer W. Mark Lanier said before the trial, “I win the case whether it is a heart attack or arrhythmia. Merck internally knew that Vioxx caused both. Merck knew it before it sold its first pill. I don’t have to use outside experts. I will use Merck’s own documents, their own emails, their own scientists.”29

Indeed, the plaintiff focused on Merck’s advertising tactics through its witnesses. Merck’s head epidemiologist on the Vioxx project responded to questions about sales techniques and pressure tactics, not epidemiological results.30 The pathologist who had written the autopsy report testified that even though there was no evidence in the autopsy of a blood clot and a heart attack, there was still a possibility that they had existed.31 That testimony was not based in scientific evidence. The plaintiffs seemed to reject the requirement for scientific evidence altogether, as demonstrated by counsel’s statement dismissing the issue that Ernst had only taken Vioxx for seven months – “That’s just bogus. Vioxx can kill you after 18 months; it can kill you after six weeks.”32

The jurors seemed to agree. They awarded Ernst $253.4 million, $229 million of which was punitive damages. The plaintiffs had asked for $40 million.33 The forensic of the jury said that the decision was about “accountability,” a response to perceptions of Merck’s actions – whether it knowingly marketed a dangerous drug – not an evaluation of the evidence for causation.34

Implications for Appeal: A Shift in Merck’s Response

According to a recent Wall Street estimate reported after the Ernst verdict, Merck’s liability for Vioxx lawsuits could eventually reach $50 billion.35 That was a sharp increase from estimates before the Texas verdict, which ranged from $8 to $25 billion.36 On the day the appeal was announced, Merck’s General Counsel Kenneth C. Frazier said that the “jury was allowed to hear testimony that was not based on reliable science and that was irrelevant.”37 Another member of the legal team said that there “is no scientific evidence that shows Vioxx causes cardiac arrhythmias,” and that the plaintiff “did not meet the standard set by Texas law to prove Vioxx caused Mr. Ernst’s death.”38 On appeal, Merck appears poised to wage exactly the sort of evidentiary attack that legal scholars have criticized as the downfall of toxic tort liability.

However, an additional defense tactic seems to be emerging. According to recent reports, attorneys for Merck contacted jurors from the Ernst trial to assess how their arguments had been received.39 Plaintiffs’ attorneys have pursued impressions and advice from the jurors more thoroughly, with paid appearances at meetings and teleconferences.40 Thus, despite the fact that the Ernst decision is not binding precedent for cases in other states, the case may have a significant legacy in future trial tactics.

News reports from the second Vioxx trial, being argued in Atlantic City, New Jersey, shed light on the lessons learned from the Ernst case. Instead of trying to limit liability through arguing causation and challenging evidence, the legal team has adopted a strategy based on what can be called “the grandmother test” – whether Merck’s leadership thought it was safe enough for their loved ones.41 Merck has attempted to introduce evidence of Vioxx use by high level employees and their families, including the current and former heads of research and the CEO’s wife.42 Such evidence has been excluded from the New Jersey case, though, on the grounds that it is not scientific.43

This issue is likely to arise in almost all of the Vioxx cases. Plaintiffs will likely try to exclude evidence of Vioxx use by Merck employees as unscientific and prejudicial because it contradicts the idea that Merck is an unfeeling corporation that rushed the product to market despite knowing it was dangerous. If the plaintiffs continue to make the intentions of Merck’s executives central to the question of liability though, as they seem to have been in the Ernst case, the decision of a judge to exclude such rebuttal evidence may be grounds for a successful appeal.

References

1. The proof required to establish liability for physical harm is outlined in Section 6 of the Restatement (Third) of Torts, a national compilation of trends in tort decisions.
2. The Doctrine of Res Ipsa Loquitur generally applies when the occurrence of an injury is enough to allow a reasonable jury inference of both negligence and causation. Rest. (2d) of Torts §928(d), on the origins of the doctrine see especially comment (a).
3. S. Pagano, “Texas Jury Awards $253 Mil-
7. Pagano, supra note 3.
8. Donald, supra note 5.
9. This limitation was created by Texas House Bill 4, a medical malpractice reform bill passed in 2003.
10. Ibid.
11. Donald, supra note 6.
12. Ibid.
13. As quoted in Donald, supra note 6.
15. Donald, supra note 6.
16. These specific proof requirements are described in Section 28, comment (c) of Restatement (Third) of Torts, the National Precedent compilation referred to in supra note 1.
20. Justice Blackmun addresses this concern in the Daubert decision, suggesting that cross examination and introduction of contrary evidence by the other side will prevent erroneous scientific testimony just as they do for other kinds of testimony. Daubert v. Merrell Dow Pharmaceuticals, Inc., 113 S. Ct. at 2797.
22. Where statistical analysis is being performed and a mean and standard deviation are calculated, the 95% confidence interval can be constructed as a range from the mean minus the standard deviation to the mean plus the standard deviation. This means that there is 95% confidence that a value will be in that range.
23. The general civil liability standard is described in Section 28, comment (b) of the Restatement (Third) of Torts, the compilation described supra, note 1. Different measures are used for different types of scientific studies (such as relative risk for epidemiological studies), and different states may apply different standards.
24. Berger, supra note 17, at 2122.
25. Wagner, supra note 18, at 774.
26. Berger, supra note 17, at 2117.
27. Berger, supra note 17 and Wagner, supra note 18.
28. As quoted in Donald, supra note 6.
29. Ibid.
30. Ibid.
31. Donald, supra note 5.
32. Ibid.
33. Pagano, supra note 3.
36. Estimate by analyst Chris Shibutani, CNN Money, supra note 34.
37. As reported in Pagano, supra note 3.
38. Ibid.
40. Ibid.
41. Ginsburg, supra note 4.
42. Ibid.
43. Ibid.
Independent Articles


Independent Articles


Paul Litton and Franklin G. Miller, “A Normative Justification for Distinguishing the Ethics of Clinical Research from the Ethics of Medical Care,” pp. 566-574.


Columns


Margaret McDermott, “Reviews in Medical Ethics,” pp. 608-610.


Staff of the American Journal of Law and Medicine, “Recent Case Developments in Health Law,” pp. 616-627.

Winter 2005, 33-4
Legislating and Litigating Health Care Rights Around the World

Colleen M. Flood, Lawrance O. Gostin and Lance Gable, ed.

Symposium Articles


Pain Management in the Emergency Department: Current Landscape and Agenda for Research

Sandra H. Johnson, ed.

Symposium Articles


Independent Articles

Paul Lombardo, “Phantom Tumors and Hysterical Women: Revising Our View of the Schloendorff Case,” pp. 791-802.


Giles Scofield, “Motion(less) in Limine,” pp. 821-833.


Columns


Tia Powell and Barbara A. Noah, “Reviews in Medical Ethics,” pp. 857-865.


Index, pp. 870-871.
Calendar of Events

DNA Fingerprinting and Civil Liberties
Date: May 11-13, 2006
Boston Park Plaza Hotel
Boston, Massachusetts

DNA Fingerprinting and Civil Liberties National Symposium represents the culmination of a three-year NIH-funded project that has brought together experts from a variety of disciplines and perspectives to study the impact of forensic DNA collection, databanking, and use on civil liberties, and to delineate areas of consensus and controversy with regard to this important issue. The results of the three-year project will be disseminated during the Symposium, along with presentations by a faculty of outstanding experts, many of whom have worked with ASLME for the past three years.

The National Symposium will provide a forum for a broad range of issues regarding forensic DNA. These issues include:

- DNA fingerprinting and civil liberties: Ethics, law and policy
- DNA in the courtroom
- Post-conviction DNA experience
- Expanding the DNA databank through familial searching
- Secondary uses of forensic DNA samples
- DNA Databanks: Who should be included?
- DNA Databanks: Issues related to the retention of samples

The American Society of Law, Medicine & Ethics will sponsor this 2 day conference May 11-13, 2006 in Boston, MA. For more information regarding the DNA Fingerprinting & Civil Liberties project please visit www.aslme.org or call 617-262-4990.

30th Annual Health Law Teachers Conference
Date: June 1-3, 2006
University of Maryland School of Law
Baltimore, Maryland

The University of Maryland School of Law will co-sponsor this two day conference intended for professionals who teach law or bioethics in schools of law, medicine, public health, health care administration, pharmacy, nursing, and dentistry. The program is designed to provide participants with updates on issues at the forefront of law and medicine and to provide them with the opportunity to share strategies, ideas, and materials.

For more information on any ASLME event, please visit our web site at www.aslme.org.
This issue of the *Journal of Law, Medicine & Ethics* is one of beginnings and endings. In the latter sense, it is the final issue that will be published with our current printer. Beginning with the Spring issue of 2006, *JLME* will be co-published by Blackwell. This move will allow our journal to be viewed by more readers (especially online) and provide greater accessibility to our journal both here in the United States and around the world. Our readers should expect virtually no change in the physical appearance of the *Journal* and no interruption of our publishing schedule. Most importantly of all, our commitment to bringing our readers the highest quality scholarship in Law, Medicine, and Ethics will go unchanged.

This edition of *JLME* also marks the first time we publish two symposiums in a single issue. This is a mark both of the high demand that the *Journal* holds with our authors and guest editors and also the great prestige of those who write for us.

One of our symposiums is guest-edited by Sandra H. Johnson, and explores the important issue of the treatment of pain in the Emergency Department. Many of the scholars in this symposium agree that pain is probably under-treated in Emergency Rooms across the world; questions arise, however, when the authors consider the related issue of prescription drug abuse, and how to combat this while still treating the most serious problem of pain. The guest editor, the authors, and ASLME all gratefully acknowledge the support of the Mayday Fund in the development of these important articles.

Our other symposium is guest-edited by old friends Lawrence Gostin, Colleen Flood, and Lance Gable, and concerns health care rights around the world. The comparative approach to studying health care has, in the past, suggested new ways of thinking about persistent medical and legal issues, as well as asked intriguing new questions. This symposium promises more of the same, continuing in the line of special symposiums in *JLME* dedicated to public health.

Finally, and most sadly, this issue will be the last in which Kathleen M. Boozang will serve as Editor-in-Chief of the *Journal of Law, Medicine & Ethics*. Kathleen served during a time of great success and change at the journal, and her influence on this institution will long out-live her tenure. Everyone at the American Society of Law, Medicine & Ethics appreciates all of her hard work and wishes her all the best in the future.

Ted Hutchinson
Editor