

ETHICS, INSTITUTIONAL COMPLEXITY AND HEALTH CARE REFORM: THE STRUGGLE FOR NORMATIVE BALANCE

*Michael J. Graetz and Jerry L. Mashaw**

This essay has two messages: First, whatever the complexity of its articulation or its interest-group politics, reforming the financing and provision of American health care is ultimately a normative undertaking. Indeed, to some it takes on the quality of a moral crusade. This is not to say however, that the ethical foundations of health care reform are unitary, even coherent. Health care reform responds to a multitude of ethical demands. The ethical “goodness” of any plan will, therefore, be a question of its “balance” among those various ethical claims.

This trite observation leads to a more complex second message — one that seems often to be overlooked by those who decry the complexity of most proposals for reforming American health care finance and delivery. The more uncertain or compromised the choices among values, the more likely the scheme will exhibit high levels of institutional complexity. For complex institutional arrangements often are devices for managing conflict and uncertainty. In this sense, a health care plan can be viewed as a design for the conduct of further struggles over both the “right” and the “good.” A system that has resolved more of these struggles in a relatively straightforward way can have a simpler institutional design. More importantly, we argue, no ethically acceptable system is likely to have wholly coherent ethical commitments. Hence, institutional complexity is a necessary price for ethical acceptability.

To elaborate on this point, we begin by indicating the broad ethical bases for the Clinton proposals and by discussing the reasons why these “ethical foundations” for health care reform are simultaneously attractive and grotesque, mutually contradictory, but also mutually supportive. Our message here is straightforward: We could choose an administratively or institutionally simpler model only if we were prepared to make certain ethical commitments now to a single ethical perspective. To the extent

* Michael J. Graetz is Justus S. Hotchkiss Professor of Law and Jerry L. Mashaw is Sterling Professor of Law at Yale Law School.

that we are not — that is, in order to hedge our ethical bets and to take a more ethically balanced approach — we must construct a much more complex regime and live with the consequences of that complexity.

Beyond this message, we have one further argument. The normative ambiguity or conflict surrounding health care reform, and the probable institutional complexity of the resulting proposals, should not blind us to the urgency of the task. In our view, the desirability of particular health care arrangements should be viewed from the perspective of the balance that a regime strikes among competing ethical commitments. No single balance is necessarily “best” in any demonstrable sense. But that does not mean that any balance is as good as any other. As an example of this latter point, our concluding section sketches our reasons for believing that America’s current health care arrangements fail virtually all the ethical tests that might be applied to a health care system, including the test of appropriate balance among ethical commitments.

I. FOUNDATIONAL AMBIVALENCE

In a refreshing departure from normal political behavior, the first draft made public of the Clinton health care reform proposals begins with a section entitled “Ethical Foundations of Health Reform.” The ethical principals enunciated, however, look in quite different directions.¹ The plan speaks of universal access to comprehensive benefits and of assuring equality of care for all Americans. These sentiments suggest that health care is a personal right with a strong egalitarian flavor. Yet, the plan also lists consumer choice and personal responsibility as essential elements of its ethical foundations, along with professional integrity as well as professional responsibility to provide effective, high quality care. These concerns reflect autonomy values often associated with individualism and markets, while also emphasizing the inviolability of the physician-patient relationship. Egalitarianism and individual autonomy are values that tend to compete in any system of social organization, at least as those values are conventionally understood.

Moreover, in setting forth its ethical foundations, the plan also talks about costs. It demands that there be a “wise allocation of resources” and a fair distribution of costs both within and across generations. Partic-

1. Of the fourteen “values and principles” set forth as the “Ethical Foundations of Health Reform,” we do not discuss three: (1) effective management; (2) fair procedures; and (3) local responsibility — the question of allocation of powers and responsibilities among national and state and local governments. On the latter, see Jerry L. Mashaw, *The Case for State-Led Reform*, 2 *DOMESTIC AFFAIRS* 1 (1993).

ularly when exhorting us to “balance prudently” what we spend on health care against “other important national priorities,” the plan speaks the language of utilitarian balancing of social needs. Once again, tension is created between this new value and those previously articulated. Making marginal trade-offs to maximize overall welfare is a language that usually eschews talk of moral or legal entitlement and focuses on social welfare rather than individual rights, personal responsibilities or professional autonomy.

In short, the Clinton health care reform plan has been grounded explicitly in multiple normative foundations that compromise the very coherence of its ethical commitments. But this compromise — indeed conflict — should not necessarily be taken to signal weakness. No single ethical perspective is likely to provide an acceptable solution to either of the two critical issues that confront any health care system: the determination of the total proportion of national resources to be expended on health care and the allocation of health care resources within that global resource commitment. To see why, we need to unpack the Clinton Plan’s ethical ideals. In each case we will do so by analyzing these issues in terms of the clusters that inhabit the Clinton health plan’s multiple ethical foundations. We will characterize these general commitments as commitments to the ideas (1) of entitlement, (2) of choice, and (3) of trade-offs to achieve the greatest good for the greatest number.

II. DETERMINING TOTAL RESOURCES ALLOCATED TO HEALTH

The overall health care budget could be determined simply by asking how much it would cost to provide everyone with all beneficial medical care. The answer seems straightforward. Since we are talking not just about lifesaving procedures, but rather of all those things that alleviate pain and increase functional capacity, there probably is no limit to the potential expenditures on marginally beneficial care. “All beneficial care” can absorb 100 percent of gross domestic product. This absolutist approach to the medical care entitlement is so self-evidently unworkable that we need not tarry in discussing it.

The practical necessity of rejecting an absolutist approach nevertheless reveals a potentially serious ethical conflict: to do so clashes with the idea of professional autonomy. For the provision of all beneficial care is the underlying ethic that now dominates the training of medical professionals. Hence, unless we wish to change medical ethics — and it is not at all clear that we do — medical providers must be induced to practice

within some set of overall arrangements that constrain the fulfillment of their professional ethical commitments.

Providing the level of care for Americans that optimizes social welfare is an obvious alternative candidate for determining the share of national resources that should be devoted to medical care. Conceptually, this utilitarian approach is simple: Spend on medical care up to the point that social spending on something else would produce greater marginal returns. In practice, of course, the computation required to achieve this goal is quite impossible. As is customary with general utilitarian calculations, we run up against two insurmountable difficulties. First, we do not know how to measure the value of various social outputs in the same metric. A bit more cold prevention versus a bit more laughter? Who knows how to do such a calculation individually, much less socially? Second, utilitarianism has never managed to overcome the principal objection of its longstanding critics — the necessity of making interpersonal utility comparisons.

Moreover, utilitarianism has the vices of its virtues. As our colleague Calabresi (who is celebrated in this symposium) has taught us, we may well not wish to allocate resources to medical care — particularly lifesaving medical care — on the basis of some overall social calculation. There is symbolic importance to affirming the value of individual lives that may lead us to want to behave in a quasi-absolutist fashion, at least in some contexts, whatever the overall social calculation of costs and benefits. The standard utilitarian response — that these symbolic values should simply be counted in the utilitarian calculus — seems to miss the essential point, for it merely circles us back to the measurement obstacles we have just described. Given these difficulties, the utilitarian approach is often simultaneously unworkable and unacceptable.

Focusing instead on the moral value of each individual life leads rather directly to the third general ethical foundation set out in the Clinton proposal, an emphasis on individual autonomy. A health care system based on the overarching value of autonomy or autonomous choice would be one prepared to allocate just that amount of national resources to health care that equals the sum that individuals would choose to spend on health care at market prices. If incomes were equal, consumer choice well-informed, and no market distortions affected the supply side of medical care, autonomous choice through markets might well be the ideal means for determining the overall national resources allocated to health care. We know, of course, that these conditions do not hold. Indeed, in the world of medical economics, they are largely absent. Moral hazard, ad-

verse selection, supply-induced demand, barriers to entry, information asymmetries and the like combine forcefully to undermine the free market program.

Not to belabor the point, the choice of a single ethical principle to guide global budget-setting in a health care system would lead fairly quickly to the unworkability or moral bankruptcy of that system. No health care system, therefore, chooses a unitary foundational principle to guide public policy-making. As with the Clinton plan and its competitors, competitive ethical principles must be combined in some complicated way, often ways that are only implicit in the construction of the plan itself. Indeed, rather than starting with ethical first principles, most health care planning — both here and abroad — is grounded in some historical baseline with respect to overall costs (such as what we are spending now) and some vague definition of the culturally accepted entitlement (i.e., “all reasonable and necessary” care).² The question then becomes how acceptable the population finds the status quo, including the means by which existing arrangements might be continued and reinforced, by comparison with the ends and means articulated in one or another reform proposal.

Policy choice then moves incrementally from a given position and within a given context, not by *tabula rasa* theorizing from ethical first principles to implementing policies. Establishing principles of entitlement, necessary trade offs with other priorities (through the inevitable balancing of uncertain costs against non-measurable benefits) and the preservation of individual choice and professional autonomy are not unimportant in this process. Indeed, they may be crucial. But they are crucial as motivations or constraints that shape arguments concerning the legitimacy or desirability of both existing arrangements and proposed changes, not as building blocks in some logico-deductive system of health care planning. Thus, if trade-offs among these ethical perspectives were not enough to induce institutional complexity, the added burdens of history and of cultural and political contingency will virtually assure that simplicity is an early casualty in all health care planning wars.

III. SETTING PRIORITIES WITHIN THE OVERALL BUDGET

Standard means of priority setting within health care systems or health care reform plans are characterized by a similar “anti-foundationalist”

2. “Reasonable and necessary” is the Canadian formulation. Borrowing from extant health insurance contracts, the Clinton plan relies on “necessary and appropriate.”

ethical approach. Priorities are not constructed from the bottom up (or is it “top down”?) by establishing ethical foundations and then taking them to their logical conclusions in the implementation of policies. To do so is to end up with an unacceptable, perhaps grotesque, system. Consider once again the usual alternatives.

Entitlement absolutists, of course, reject the necessity of setting (explicit) priorities. But, because the failure to make choices is the health care equivalent of rejecting the force of gravity, absolute entitlements approaches must smuggle in implicit priority-setting devices. One possibility might be to define strictly what is and what is not medical care — identifying things that are in and things that are out of the medical care system. But, since the absolute entitlement approach has no internal criterion for denying the appellation “medical care” to any product, procedure or social action that potentially would improve individual or population health, the explicit rejection of anything arguably beneficial to health is necessarily, and all too obviously, arbitrary.

An alternative and less visible strategy relies on a delegation of authority to the medical profession to define “necessary or appropriate care” in the context of practice. Fear of allocating all of GDP to health care, however, has led other nations to bound this discretion within some budget constraint thereby denying in the aggregate what they proclaim at the level of individual practice. This may be a politically saleable position, but it clearly conflicts with the ethical posture of absolute entitlements.

Nor need we extensively critique the autonomous market-choice solution to prioritizing within the health care budget. Certainly, a “free market” is a clear and effective means of priority determination. Indeed, the market sets priorities and the overall budget simultaneously as the simple vector-sum of spending by individuals. But, a badly distorted health care market is a no more reliable mechanism for implementing autonomy in priority setting than in budget building. Market defects, such as moral hazard and suppliers’ control over demand, distort decisions at the individual level as well as in the aggregate. This is not to say that using certain “market like” devices and incentives to force people to face up to some of the costs of their choices might not be a good idea. Our point is merely that an autonomy-through-markets solution is not a general solution to priority setting *within* health care any more than to determining overall budget allocations *to* health care.

Although the measurement problems remain problematical, certain variations on the utilitarian theme seem to have greater attraction in the

priorities arena than they had when attempting to establish the size of the overall budget. There is, after all, the possibility that, with the budget being set, some rough and ready calculations are possible with regard to what the most effective use of those funds might be within the domain of health care. However, once again, there are some serious problems. Cost effectiveness analysis must ask "cost effectiveness for what?" Additionally, in the arena of health, the specification of the "what" may produce radically different, and often objectionable, allocations of resources within the health care budget. We will discuss just three possibilities to provide a flavor of the difficulties:

One approach was pioneered on the battlefield and marches under the label "triage." In its battlefield form the "Cost effective for what?" question is answered: "For life saving, stupid." Persons generally are sorted into three groups: Those who will live whether or not they receive immediate medical care; those who will die whether or not they receive immediate medical care; and those for whom immediate medical care will make the difference between living and dying. Giving immediate care to the third group is the clear priority.

While the implicit definition of health care as only that which is life saving may make perfect sense under battlefield conditions of extreme scarcity, it is not acceptable as an overall solution to setting priorities within the health care system. Once we leave the battlefield, the preservation of life is not the only thing we care about. Therefore, spending our whole health care budget on preserving life, whatever its quality and however long or short, is not satisfactory as a simple rule of decision. In many instances, we will look not only for other measures that include the alleviation of pain, the improvement of functioning, and a host of other values when determining what our medical care system is for, but we will also find many circumstances for elevating these over heroic life saving measures.³

An alternative approach would be to allocate medical care resources in ways that maximize the restoration of normal functioning for the whole of the population. This need-based approach to cost effectiveness in medical care is both attractive and grotesque. Taken to its logical conclusion, it would cause us to spend most, if not all, of our resources on those who had the worst medical conditions and who needed the greatest care

3. Triage may still be used, but for different and far more limited purposes, such as who gets seen in a walk-in or urgent care clinic.

and assistance in order to achieve what was agreed to be the equivalent of "normal" social functioning.

While there is much to be said for greater expenditures on those who are "least advantaged" in Rawlsian terms, it is hard to believe that we wish to prioritize all our medical resources in this way. Indeed, assuming that we are operating within a budget constraint that binds, an attempt to structure priorities to maximize normal functioning for each individual would require eliminating virtually all interventions that merely alleviate pain or speed up the recovery of people who would approach or return to normal functioning without any expenditure of social resources on their medical care.

"Allocate all resources to restore or create normal functioning for each individual" is, thus, an inadequate unitary guide to health care priority setting. It implicitly abandons both the relatively well and the dying in favor of a strong, perhaps even exclusive, allocation of health care resources to those with serious conditions, but some potential for normal functioning if given heroic allocations of medical care services and technologies. Like an exclusive focus on lifesaving, this approach is far too narrow to satisfy our moral demands for a more "balanced" approach to priority setting.

In an effort to add more scientific precision and balance to cost effectiveness analysis in medical care, some have suggested using the idea of a "quality adjusted life year" (QALY) as the metric for comparing the value of medical care expenditures.⁴ The basic idea is just this: A year of perfect health forms one end of the continuum of quality, death the other. A year of perfect health is rated at one and death at zero. Any impairment or illness that detracts from perfect health reduces an individual life-year from a value of one to something between one and zero. Hence, the life year has been "quality adjusted" for health losses.

These "quality adjustments" are done by panel surveys in which respondents are asked to determine how many years of life with certain illnesses or disabilities they would be willing to trade for a year of perfect health. If respondents who have — or, more usually, who are asked to imagine they have — condition A would trade twenty years of life with that condition for ten years of life with no health problems, then condition A gets a rating of .5 in the quality adjustment scale. Any medical

4. For a more extensive analysis, see Chapter 5 of PAUL MENZEL, *STRONG MEDICINE: THE ETHICAL RATIONING OF HEALTH CARE* (1990). See also Alan Maynard, *Logic in Medicine: An Economic Perspective*, 295 *BRITISH MED. J.* 1537; John Harris, *QALYing the Value of Life*, 13 *J. MED. ETHICS* 117 (1987).

intervention that will alleviate condition A can be determined to be relatively effective or ineffective by a computation which compares its cost with the cost of some other procedure that alleviates a similarly rated condition for someone having a similar number of years of life before them. Thus, for example, if conditions A and B both create .5 valued quality adjusted life years, but the costs of alleviating B are less than the costs of alleviating condition A, the procedures that alleviate B will have a higher priority in a health care system that measures cost effectiveness in terms of quality adjusted life years.

The QALY approach may appear precise and even sensible, but it has some rather obvious problems, not the least of which is its capacity to produce ethically grotesque results. For example, the alleviation of a condition (x) which is quite trivial in terms of the loss of quality of life (say .01) may nevertheless be cost effective if it is extremely inexpensive. After all, if curing condition x costs only five cents, the expenditure of \$25 could buy 100 such procedures for 50 persons having a one year life expectancy producing a QALY value of 5. However, using this sort of calculation implies that such an expenditure would be preferred to the expenditure of \$25 on saving a single individual from an immediately life threatening condition, where the person in question has a four year life expectancy. Five QALYs beats four QALYs, even if this literally means the provision of 100 bandaids and no tourniquets. The ease with which similar examples can be generated is alone sufficient to make QALY analysis unattractive as the singular guide to the expenditure of medical resources.

But this is not the only feature of the QALY approach that may be morally objectionable. Because evaluating QALYs entails a consideration of the number of remaining years of life, and not just their quality, there is a certain "ageism" built into the procedure. To be sure, many persons are reasonably comfortable with preferring the young to the aged in situations of dire necessity. But, a similar advantage with regard to routine allocation of medical care is clearly less morally attractive.

Moreover, this sort of distinction between persons based on one aspect of their individuality — in this case, age — highlights the basic underlying disquiet that the QALY system induces in most persons. Do we as a society really want to make allocation decisions based on gross indicators such as age, or the *presumed* quality effects of some health condition on a particular individual's enjoyment of life? Are we willing to accept the systematic disfavoring in such a system of anyone with a chronic illness or disease and for whom medical intervention for acute and life threatening

conditions could never have a value of one times life expectancy? Does not the QALY procedure simply mask our disquiet about poignant trade-offs among individuals, all of whom have unique but differentiated claims to social resources, through a scientific quantitative formula that ignores the human dimensions that are necessary to humane decisionmaking about medical care priorities?

Once again, however, to reject the QALY approach as the basic guide to setting priorities within a health care system is not to say that it has no value for any purpose. When the QALY methodology is used to compare resource allocations across health care systems, it may reveal striking differences in medical care allocation decisions and challenge us to think carefully about the allocations made within our own system. Yet, it is quite unlikely that many of us could come to believe that we ought to think about health care systems in a world series or world cup format in which the system with the best QALY scores was for that reason alone the most ethically attractive health care system.

Given the sensible ethical objections to each of the absolutist, market-based and cost-effectiveness solutions to internal priority setting within medical care systems, it is hardly surprising that we see mixed approaches to the determination of medical care priorities. As in the case of global budget setting, nations "muddle through," using some set of historically contingent coverage definitions and institutional arrangements to economize on utilization. Notions of entitlements, market incentives and cost effectiveness evaluation cohabit uneasily in setting the priorities of any medical care system. Yet, cohabit they must, for we are understandably unwilling either to use any single method alone or to reject any one of them as entirely without merit.

How then should we think about health care systems from an ethical point of view? If the idea is not to determine basic ethical foundations both for building overall budgets and for prioritizing within them, and then developing policies to implement those ethical commitments, how are we to proceed? Our basic, admittedly untidy, claim is that we do and must approach these issues from an essentially Aristotelian perspective, asking: What is the *balance* of normative commitments structured into the system and how does this balance comport with our sense of appropriate total resource allocation and internal priority setting? This is an untidy approach because it attempts to incorporate all of the ethical foundations that are moderately attractive in making resource allocation decisions, while not giving a clear preference to any one of them. Moreover, it asks a question which is highly contingent and culturally contextual.

The appropriate “balance” of the system will be perceived differently by different observers and may vary widely across cultures or even within the same culture at different times. Indeed, the “approach” that we suggest here as necessary in health care arrangements may be so complex, ambivalent and vague that it has little critical bite for just the issues that must be confronted.

To some degree this criticism is fair. Yet, we maintain that there is enough here to help explain, in ethical terms, why certain solutions tend to be favored at this time in our society above others. In particular it can explain why health care reform is currently such an urgent priority for many Americans.

IV. THE ETHICAL INADEQUACY OF CURRENT ARRANGEMENTS

One straightforward explanation for why health care reform is currently high on the agenda of American politics is that our current arrangements fail *all* reasonable tests of ethical institutional design. The entitlements approach is clearly not satisfied by a system that leaves thirty-five to forty million persons (many of them children) uncovered, and many more fearful of inadequate or lost coverage. On the other hand, for those who enjoy adequate health insurance coverage, the current system seems to incorporate an absolutist prejudice that threatens us with national bankruptcy. The persistence of supply-driven demand, a third-party-payer system that makes costs invisible at point of service, and a tax system that fosters substitution of health insurance for cash wages and consumption of health care (particularly by the affluent) in preference to other goods, can hardly be taking the maximization of general welfare — or even cost-effectiveness — very seriously.

Moreover, as we discussed above, we know that the health care market is structured in such a way that the underlying ethic of markets, their responsiveness to autonomous individual choice, is virtually absent from American medical care. Our “choices” reflect the judgments of uninformed patients propelled by the supply side inducement of ubiquitously available high-tech high-cost diagnostic and treatment methodologies and the absolutist professional tendencies of physician “advisers.” “Patient autonomy” under current arrangements is thus often mocked by dual visions: one is of the well-insured patient trapped in a high-tech bondage that fails to minister to her most basic human needs; the other is of the uninsured, who must “choose” impoverishment to obtain adequate care.

In short, our current health care system balances our ethical demands

very poorly indeed. It may well be possible to construct a system in worse ethical balance than the one that we have. Perhaps we will do so. But, contemplation of the status quo makes most who reflect upon it urge taking the risks of reform, however complicated and compromised, rather than accepting the continuation of the current, equally complex, but morally bankrupt, system.