

Health Maintenance Insurance: Toward an Optimal HMO*

KENNETH E. WARNER

Department of Health Planning and Administration, School of Public Health, University of Michigan, Ann Arbor, Michigan 48109

ABSTRACT

The financial incentive structure of today's health maintenance organizations addresses certain problems attributed to fee-for-service medical care, but at a theoretical level it does not induce optimal provider behavior. Health maintenance insurance—a combined package of medical, morbidity/disability, and life insurance—encourages providers to compete for the *health* dollar, and not simply the *medical care* dollar, thereby remedying deficiencies in prepayment and promoting true health maintenance. The principle underlying health maintenance insurance emphasizes the need to search for effective means of preventing disability, morbidity, and premature death.

Introduction

Prepaid comprehensive health care, as embodied in the Health Maintenance Organization (HMO), is often advocated as a solution to deficiencies in the delivery of personal medical care and as a means of bringing the spiraling costs of care under control. Proponents argue that fee-for-service, the predominant mode of payment, encourages or at minimum tolerates incorrect and often excessive use of medical resources. By altering the economic incentive structure and making providers responsible for their patients' long-term medical needs, prepayment is intended to introduce economic efficiency into the delivery of effective health services (e.g. Ellwood et al., 1971).

Long the fair-haired child of the academic health community, prepaid comprehensive care entered the federal policy arena with debate which culminated in passage of the HMO Act of 1973. HMOs were offered to Congress as a competitive market alternative to growing regulation in the health field, but the regulatory maze written into the legislation and the paucity of start-up appropriations emasculated the HMO movement (Starr, 1976). Recent amendments to the act have not significantly reversed the damage, but attraction to the idea has reemerged in the policy debate on national health insurance (Enthoven, 1977).

Conceptually, prepayment successfully addresses many of the weaknesses of fee-for-service; and recent empirical evidence indicates that these conceptual advantages

* This is a revised version of a paper presented at the 105th Annual Meeting of the American Public Health Association, Washington, D.C., October 30–November 3, 1977.

have frequently translated into real benefits in practice (e.g. Roemer and Shonick, 1973; Weil, 1976; Luft, 1978a). However, even in theory prepayment does not produce the set of economic incentives which would be optimally structured to produce health maintenance. The purpose of this paper is to note the deficiencies remaining in the concept of prepaid comprehensive care, as regards economic considerations, and to suggest a simple modification which would create a theoretically complete set of economic incentives to encourage true health maintenance.

A couple of points merit emphasis at the outset. First, I am concerned here exclusively with the economic environment in which care is demanded and delivered. I do not mean to imply that economic factors are the only factors influencing medical decisions, nor that they are the most important. Economic considerations constitute only one element in a constellation of influences including technical, social, ethical, and other professional considerations. My concern is that the economic environment be such as to support and encourage outcomes deemed socially desirable. Thus, this analysis focuses on the *marginal* effects of economic incentives—the incremental effects one would anticipate *theoretically* when all other factors are held constant.

Second, I beg the reader's indulgence of a few generalizations made in the following pages. Their use seemed consistent with the objective of conveying the principal idea simply and concisely. The assumptions underlying the concept of health maintenance insurance and its implementation can be reasonably discussed only on a foundation of shared basic understanding of the concept itself.

Before introducing that concept, I will briefly review the economic incentives implicit in both fee-for-service and prepayment.¹

Economic Incentives in Fee-for-Service

The primary argument in favor of fee-for-service medicine is a conventional economic market argument: out-of-pocket charges deter consumers from demanding care which they value less than the resources used in providing that care. In this view, the absence of direct out-of-pocket financial responsibility encourages overutilization (wastage) of scarce resources. Critics of fee-for-service respond that consumers may be ignorant and myopic about their health, and certainly about specific medical care needs. Out-of-pocket charges may exacerbate a tendency to underutilize primary and preventive care services, resulting in serious—and more expensive—medical problems later on. These critics argue that barriers to the consumer's initiation of demand, such as out-of-pocket fees, ought to be eliminated, with the restriction of demand being accomplished by a system of triage in which medical professionals determine in each instance whether or not, what type, and how much medical care each patient should receive. The force of both the pro and con sides of this debate is mitigated somewhat by the fact of extensive medical

¹ These have been discussed frequently in the literature. An excellent early discussion is that of Havighurst (1970). Luft (1978b) presents an interesting analysis of the effects of these incentives on use of preventive services.

insurance coverage and the resultant increases in demand due to the lack of out-of-pocket financial responsibility for insured services.

The incentives in fee-for-service can influence providers as well as consumers. Providers have few disincentives, and may have positive incentives, to overprescribe medical goods and services (e.g. Schroeder and Showstack, 1978). Of critical importance is the fact that as agents for their patients' demand, physicians make consumption decisions, yet usually they bear no financial responsibility for resultant use of goods and services. That burden is borne by the patients and by third-party insurers. Insofar as increased consumption translates into increased income, providers have a positive economic incentive to encourage consumption. Presumably the patient cannot distinguish needed consumption from that which is largely income-augmenting for the provider. To the extent that patients are well insured, they may not care.

A tendency towards excessive use of resources by physicians is not motivated exclusively nor even necessarily primarily by greed. The desire to deliver "Cadillac care" is partly a response to the physician-patient trust relationship. "[T]he social obligation for best practice is part of the commodity the physician sells The safest course to take to avoid not being a true agent is to give the socially prescribed 'best' treatment of the day. Compromise in quality, even for the purpose of saving the patient money, is to risk an imputation of failure to live up to the social bond" (Arrow, 1963). Medical ethics may encourage excessive use of resources; the economic environment may permit it.

"Neutrality" of fee-for-service for the physician is seen in the growing practice of defensive medicine. Such procedures as skull X-rays following a head injury are routinely performed by physicians who recognize their limited medical value but order them to cover themselves against malpractice charges (Office of Technology Assessment, U.S. Congress, 1978). When the costs of such procedures can be easily passed on, there is no economic incentive to restrict resource consumption to procedures of demonstrable value.

An indirect fee-for-service incentive effect relates to the predominant mode of insurance: most present-day insurance provides fuller coverage for hospitalized patient care than it does for ambulatory care. This is based on the belief that the sorts of care one would seek in an ambulatory setting are likely to be more discretionary—that is, more sensitive to price—than care for which hospitalization is appropriate. Thus, the structure of insurance provides an incentive for patients to be hospitalized rather than be treated in the cheaper ambulatory setting, which would often be at least as adequate. This is generally assumed to have had a deleterious effect on the containment of medical care costs.

Finally, consider how the system of fee-for-service relates to the supply of and demand for health education. It is clear that, while personal habits and life style are of fundamental importance in determining health, millions of people engage in potentially self-destructive habits. For many, such behavior is knowingly undertaken; for these people, the consumption value of the habit may exceed its dangers. For others, the habits represent ignorance. In either case, one must ask whether or not there is a significant demand for health education. Certainly there is

little push to supply health education activities commercially. At the theoretical level, the medical profession actually confronts a disincentive to promoting healthful habits and life styles: to the extent that it succeeds, it reduces the number of patients presenting themselves with medical complaints. Thus the economic factor may encourage the treatment of illness rather than the promotion of health.

Economic Incentives in Prepayment

Advocates of prepayment claim that it addresses both sets of undesirable incentives associated with fee-for-service: those affecting consumers and those affecting providers.

For consumers, prepayment eliminates or substantially reduces the out-of-pocket financial barrier to seeking care (the latter if a small visit co-payment is charged or if transportation or other related costs are a factor). As noted above, some observers, especially many economists, fear that this sows the seeds of overutilization by consumers. Prepayment advocates respond that providers are best equipped to determine what constitutes desirable utilization. This is most easily achieved by encouraging consumers to present themselves to lower-echelon health professionals who filter them to the appropriate level of care. Overutilization is protected against by a system of triage, so that truly excessive demand is restricted to the low-cost point of entry. (There are additional means of discouraging excess demand. For example, nonpecuniary costs, such as increased waiting time, can be imposed on patients.)

For providers, the incentive associated with prepayment *by itself* is toward underutilization. Under prepayment, marginal resource consumption simply adds to the cost of operating the HMO; the expense is not passed on to the patients (though, of course, it might affect membership charges in later years). Thus, additional resource consumption simply reduces the HMO's net revenues and consequently the partners' incomes.² However, the incentive for providers to underutilize resources is offset by two economic considerations: if underutilization results in observably poor health care, the HMO will lose members and may not be able to attract new members. The organization's survival depends on its having a presentable image. Secondly, the HMO is responsible for its members' medical care throughout the period they remain members. If insufficient care in the current year results in a more serious illness or disability in later years, the HMO may be saddled with large expenses which were preventable. Thus the providers' economic incentive is to minimize discounted resource consumption expenses *over time*. That may well mean that incurring a relatively small expense today is likely to be profit-maximizing behavior for the HMO in the long run.

Similarly, the HMO has an incentive to undertake health education and to encourage primary and preventive care to the extent that such activities decrease the later incidence of those types of morbidity and disability which require expensive

² For the individual medical decision maker, this incentive diminishes as the number of partners rises, since the marginal cost of that individual's additional resource consumption is spread over larger numbers of people.

medical care. In theory, the HMO would weigh the current cost of preventive activities against the discounted value of the expected medical costs which could be averted in the future (the latter constituting the benefits of the preventive activities) and undertake activities which demonstrated positive net benefits.

Despite its many attractions, from a theoretical point of view prepayment includes some perverse economic incentives. While these incentives are checked somewhat by HMOs' members or prospective members observing the results of inadequate care, consumers' ability to observe negative health outcomes and associate them with inadequate care is restricted by their knowledge and by the time frame involved. Even if such observation is occasionally possible, the *direct* economic incentive to providers remains nonoptimal. The perverse incentives are:

(a) Prepayment discourages treatment which will have the effect of keeping someone alive who will require long-term medical care. For example, given a patient with end-stage renal disease, the prepayment incentive *by itself* (i.e. taken apart from humanitarian considerations) favors letting the patient die rather than placing the patient on renal dialysis, an expensive procedure which could keep the patient alive (but on dialysis) for years.

(b) More generally, and less dramatically, if medical care can prevent premature mortality, it adds to the future pool of the elderly for whom the providers will be responsible, and the elderly require more care than younger people. In this context, with the provider having long-term responsibility for patients' care, prepayment does not encourage the medical promotion of longevity.

(c) Prepayment discourages treatment which will prevent later morbidity or disability whose primary costs are nonmedical, i.e. costs associated with suffering or incapacity. For example, a physician's failure to set (or poor setting of) a leg bone fracture can result in an arthritic condition later in the patient's life. The chief costs associated with this condition would be both nonpecuniary (pain and possibly a limp) and pecuniary (e.g. additional housekeeping assistance) but in neither case medical. The financial incentive in prepayment encourages effective current treatment only to prevent significant later *medical* costs which the HMO will have to absorb.

(d) Similarly, prepayment discourages health education which has beneficial health effects which do not significantly decrease later medical costs and may actually add to such costs at the same time that they increase people's longevity and sense of well-being.

The intent of the proposal in the next section is to correct these undesirable incentives and strengthen the sound foundation of the prepayment concept.

A Health Maintenance Insurance Package

While no one has developed a completely satisfactory definition of health, it is clear that two major components of health are freedom from disability and morbidity, and the fact of life itself—freedom from death for as long as life can be enjoyable. Achieving these conditions constitutes health maintenance; the incentives in the HMO should encourage this. The incentives should not bias the HMO in favor of

preventing only health problems with significant associated medical expense, ignoring those that involve little medical care. Certainly there should be no positive financial incentive to let medically dependent patients die.

An organization will be encouraged to achieve true health maintenance if it is rewarded financially for producing health, not simply for avoiding medical conditions. The proposal here is to improve the HMO's incentive structure by having it offer consumers comprehensive *health maintenance* insurance (HMI), rather than the comprehensive *medical* insurance which it currently provides. This can be realized by offering an insurance package which includes:

- comprehensive medical insurance (the current HMO insurance)
- disability/morbidity insurance
- life insurance

Under such an insurance plan, the owners of the HMO would profit from their members' living long lives free from illness and disabling conditions, and not simply free of conditions requiring a great deal of medical care. Under the present system, theoretically the financial inducement is to compare the cost of a current treatment with the discounted expected cost of later treatment which would result from failure to treat today (i.e. the discounted cost of later treatment weighted by the probability that later treatment will be required due to the absence of treatment today). Current and possible future treatment are the alternatives. If the former is cheaper than the latter, economic considerations favor treating the patient today; if the relationship is the opposite, the economic incentive recommends foregoing treatment today and accepting the risk that treatment will be required in later years.

But the true health cost of this problem is not simply the medical cost. Suppose that the failure to treat a condition today does eventually result in a complication requiring medical treatment. Suppose further that associated with the complication or with the later treatment is a period of morbidity or disability during which the patient is incapacitated and/or experiences suffering. This is of no direct economic concern to current HMOs. However, if the HMOs were financially liable for the morbidity or disability, they would then include this cost in their treatment decision calculus. That is, other things being equal, physicians would be more likely to treat patients today rather than later because the true costs of the later complication are greater than the medical costs alone. Of course, the opposite situation is also possible: inclusion of disability/morbidity costs might recommend postponing treatment if current treatment was the source of disability (for example, the patient's being bedridden following an operation). Again, however, this accurately reflects consideration of all of the true costs of treatment.

Under health maintenance insurance, the decision of whether or not to administer life-maintaining therapy which may continue until the patient's death is no longer a battle between the physician's conscience and the financial realities of continued treatment. Now a decision not to administer therapy—hence to allow the patient to die—also involves immediate payment of a life insurance policy to the patient's beneficiaries. Treatment which prolongs life postpones the burden of the life insurance; if the policy is for term insurance, life-maintaining treatment decreases

the face value of the death payment as well. It is conceivable that this structure of financial responsibility would make long-term treatment the HMO's profit-maximizing strategy. Even if it does not, at least the provider's conscience may find an ally in the life insurance liability. The economic incentive is thereby tailored more closely to the humanitarian objective.

In general, the addition of life and disability/morbidity insurance to the HMO's package means that the organization has more incentive to be concerned with members' total health and not simply their medical health. Theoretically, consequences of HMI include the following:

(a) Given the client population, the HMO has a greater inducement to learn the true value of such activities as health education and primary and preventive care. The incentive to learn the link between these activities and future medical needs is strengthened, and the HMO has the additional incentive of undertaking activities which have nonmedical health rewards. In effect, health maintenance insurance would subject a number of conjectures about primary and preventive care to the rigors of a market test. (The current HMO setting does the same, though again with a medical bias. HMI strengthens the incentive to discover the relationships, since more costs are involved.)

(b) Being financially responsible for members' total health reduces the HMO's incentive to take shortcuts in the provision of useful medical care.

(c) Proponents of HMOs argue that prepayment and comprehensive care will restore competition to the medical marketplace, and that this will have the effect of encouraging greater efficiency (Ellwood et al., 1971; Enthoven, 1977). The tendency toward competition would seem to be strengthened by the inclusion of life and disability/morbidity insurance in the HMO package: the package is more comprehensive than that offered by the current HMO, hence it is also more expensive. Presumably, consumers will shop more carefully when the commodity involved is as financially significant as this package would be. Note again, however, that there are no incentives for the HMO to cut corners in the provision of care. The profitable HMO will be the efficient provider of health maintenance services. This will be the organization which can expect minimum average total costs and hence the greatest margin between revenues and expenditures.

Implementing Health Maintenance Insurance

The principle of health maintenance insurance could be put into practice in several ways. First, of course, a few HMOs might actually adopt the practice. To do so, an HMO would have to have a critical mass of membership, since insurance requires a large population to smooth out risk. Today's largest HMOs would have the minimum requisite membership. Smaller HMOs could band together for insurance purposes while continuing to provide decentralized patient care.

Big firms and labor unions could be implementing HMI right now. They are concerned with each of its components—medical, life, and disability insurance. Clearly it is in the mutual interest of management and labor to seek an efficient health care package, one which delivers the benefits desired by labor at the low cost

desired by both management and the unions. This could be achieved either by working with a single HMI provider/insurer or by contracting for medical care coverage which, because of its preventive care aspects, would induce life insurers to lower their premiums to the group. As is noted below, the limited attention devoted in contract negotiations to the interrelationships among medical, disability, and life insurances is perhaps a telling comment about the perceived connection between medical care and morbidity and mortality outcomes.

A plausible implementation of the HMI principle is through decentralized market forces. In a world of perfectly competitive insurance companies and HMOs, the former would rate the latter and would scale life and disability insurance premiums accordingly. That is, if the members of HMOs of type A experienced less morbidity and mortality than the members of HMOs of type B, members of type A HMOs would pay commensurately less for life and disability insurance. Assuming that they served similar populations, type B organizations would be forced by competitive pressures to restructure their health care activities to imitate their more successful rivals. Of course, both the insurance business and the medical care delivery business are far from the competitive ideal. Still, if truly significant differences in health status do emerge from different types of delivery organizations, several insurance companies can be expected to introduce differential premiums. Where differences in outcomes are substantial, such differences are generally ultimately reflected in premiums. Consider, for example, “good driver discounts” on auto insurance and lower life insurance rates for nonsmokers. The insurance market may adjust slowly, and medical care delivery organizations may respond at least as slowly, but as evidence accumulates on variations in morbidity and mortality, competitive market pressures can be expected to lead the system in the right direction.

Any application of the HMI principle—within a single HMO, through union benefit packages, or through insurance ratings of HMOs by independent insurance firms—would benefit from partial or complete resolution of a number of very difficult questions. For example, how does one measure morbidity and disability? How are their true costs to be assessed? (Workmen’s compensation and similar plans offer partial but not wholly satisfactory answers.) How does one evaluate relative programmatic effectiveness in the face of significant differences in the socioeconomic and demographic composition of different HMOs? (Many questions such as these are equally germane to consideration of today’s HMOs.) Further, implementation of health maintenance insurance by a single organization—an HMO providing the complete insurance package—would necessitate clarification of state insurance laws and answers to questions such as, what happens to a person’s health maintenance coverage when the person moves? Would all members of an HMO, including children, be permitted or compelled to carry all of the insurance components of HMI? And so on.

Conclusion

Health maintenance insurance is unlikely to find widespread acceptance in practice in the near future. In addition to the deterrents enumerated above, there are few

competitive pressures to bring it about. Indeed, both professional resistance and government regulation serve as powerful anticompetitive constraints on innovation in medical care delivery and financing (Havighurst, 1977). And, as regards the single-organization approach, many health professionals would consider the HMO's going (further) into business to be distasteful. However, simply enunciating the principle can be of value. The principle behind HMI points to conceptual inadequacies in the economic incentive structure of current HMOs; it emphasizes the distinction, too often ignored, between medical care and health care; and it illuminates the role that economic incentives might play in encouraging providers both to seek new methods of promoting health and to attempt to alter their members' unhealthful habits.

One might acknowledge the theoretical appeal of HMI but find its practical importance limited. The economic incentives implicit in prepayment may not be theoretically optimal, this line of thought runs, but medical and humanitarian considerations dominate economic factors when the latter deviate from the ideal; in practice, HMOs *are* concerned with the preservation and enhancement of health; their behavior is directed toward true health maintenance.

Obviously, this is an empirical question, one which cannot be adequately answered with the paucity of existing understanding of the effects of medical care and health behavior on health status. However, recent empirical evidence challenges the contention that HMOs employ preventive care procedures more often than the fee-for-service system (Gaus et al., 1976; Luft, 1978b) and theoretical work other than the present paper has been offered to explain why HMOs might underproduce preventive care (Finkle, 1975). Perhaps the most significant evidence to date is that alluded to earlier, namely the indifference of the marketplace: life insurers are not offering lower premiums to members of HMOs, suggesting that insurers do not currently perceive a significant relationship between HMO care and lifespan; and union and management contract negotiators rarely make a direct connection between medical coverage and life and disability insurance, despite the substantial economic incentives to both groups to reduce the cost of fringe benefits. To be sure, these facts might simply reflect noncompetitiveness or slowness of adjustment in these markets; alternatively, the relevant decision makers might be myopic, ill-informed, or conservative. But the fact remains that HMOs have not reduced the cost of life insurance, and labor contract negotiators have evidenced relatively little faith in the idea that medical care can be restructured to significantly combat mortality and morbidity.

It might be argued that health maintenance insurance is a more drastic departure from the standard HMO model than is necessary. Specific health education and medical practices which are demonstrably effective could be promoted by a system of direct bonuses to providers, as is currently done in Britain's National Health Service. However, the need for activity-specific bonuses points to deficiencies in the existing incentive structure. Implicit in health maintenance is a similar set of bonuses possibly covering a larger number of such activities. Furthermore, HMI identifies cost-effective activities, not simply "effective" ones, and scales the implicit bonus for each closer to its true value than does an arbitrarily established figure. Of

perhaps greater importance, HMI encourages medical managers to seek out new means of promoting health. A defined set of activity-specific bonuses offers no comparable incentive.

Both in theory and in practice, HMOs seem to represent a significant step in the direction of more efficient delivery of medical care. Unlike fee-for-service, prepayment introduces incentives for providers to be cost conscious in their prescribing and delivery of care. Empirical studies repeatedly report lower than average rates of surgery and hospitalization for HMO members (e.g. Weil, 1976; Luft, 1978a). The attractiveness of health maintenance insurance is that the substantial amounts of money involved in disability and life insurance—and hence in HMI—would place a significant new pressure on providers to search for truly effective means of preventing disability, morbidity, and premature death. It seems only logical that HMOs would develop more concern for true health maintenance if they bid for the *health* dollar, and not simply for the *medical care* dollar. Health maintenance insurance would make the economic incentive structure of the HMO more compatible with the humanitarian considerations which everyone would like to motivate the delivery of personal health care.

Acknowledgment

This paper grew out of discussions with Joseph Chadwick. He deserves much of the credit for any good ideas expressed herein, and perhaps a share of the blame for any weaknesses.

REFERENCES

- Arrow, K. (1963). "Uncertainty and the Welfare Economics of Medical Care," *American Economic Review*, Vol. 59: 941–973.
- Ellwood, P. et al. (1971). "Health Maintenance Strategy," *Medical Care*, Vol. 9: 291–298.
- Enthoven, A. (1977). "National Health Insurance (NHI)." Memorandum for DHEW Secretary Joseph Califano.
- Finkle, W. (1975). "Quality of Care Versus Cost Containment: Implication for Professional Standards Review Organizations." Paper presented at the Annual Meeting of the American Economic Association, Dallas.
- Gaus, C., Cooper, B. and Hirschman, C. (1976). "Contrasts in HMO and Fee-for-Service Performance," *Social Security Bulletin*, Vol. 39: 3–14.
- Havighurst, C. (1977). "Controlling Health Care Costs: Strengthening the Private Sector's Hand," *Journal of Health Politics, Policy and Law*, Vol. 1: 471–498.
- Havighurst, C. (1970). "Health Maintenance Organizations and the Market for Health Services," *Law and Contemporary Problems*, Vol. 35: 716–795.
- Luft, H. (1978a). "How Do Health-Maintenance Organizations Achieve Their 'Savings?'" *New England Journal of Medicine*, Vol. 298: 1336–1343.
- Luft, H. (1978b). "Why Do HMOs Seem to Provide More Health Maintenance Services?" *Health and Society*, Vol. 56: 140–168.
- Office of Technology Assessment, U.S. Congress (1978). *Assessing the Efficacy and Safety of Medical Technologies*. Washington: U.S. Government Printing Office.
- Roemer, M. and Shonick, W. (1973). "HMO Performance: The Recent Evidence," *Health and Society*, Vol. 51: 271–317.

- Schroeder, S. and Showstack, J. (1978). "Financial Incentives to Perform Medical Procedures and Laboratory Tests: Illustrative Models of Office Practice," *Medical Care*, Vol. 16: 289-298.
- Starr, P. (1976). "The Undelivered Health System," *Public Interest*, Vol. 42: 66-85.
- Weil, P. (1976). "Comparative Costs to the Medicare Program of Seven Prepaid Group Practices and Controls," *Health and Society*, Vol. 54: 339-365.