

The Performance of Health Maintenance Organizations: An Analytic Review

FREDRIC D. WOLINSKY

*Center for Health Services Research and Development,
American Medical Association*

EVER SINCE 1970 WHEN PAUL ELLWOOD (1971) FIRST coined the phrase "health maintenance organizations" (HMOs) and Richard Nixon (1971) quickly made them the official national policy goal, HMOs have been the most touted, discussed, analyzed, and hotly debated alternative health care delivery system in the United States. Coming into formal existence through the HMO Act of 1973 (PL 93-222) and its subsequent amendments (Department of Health, Education, and Welfare, 1974, 1975; PL 94-460; PL 95-559), HMOs have been officially targeted as a national priority (PL 93-641) and as one of the potentially key components of a national health insurance system. Accordingly, one would assume that by now a common understanding of what HMOs are, what they are supposed to do, and what and how well they actually do it would have emerged. Unfortunately, this is not the case. In the present paper we seek to resolve part of this problem by 1) reviewing the nine most often cited reviews of the HMO performance literature; 2) identifying the incentive and disincentive structures operating in HMOs; 3) exposing the methodological problems associated with evaluating the performance of HMOs; and 4) analytically reviewing the recent literature evaluating the performance of HMOs.

Findings and Conclusions from Previous HMO Performance Reviews

Although there has not yet been a definitive evaluation of HMO performance, the issue has been discussed in several hundred papers, articles, and books. In particular, there have been nine reviews of these evaluations (or, if you will, nine “state of the HMO field” papers), which are generally considered to be quite informative, if not authoritative. We shall begin by highlighting the findings and conclusions reached in these nine reviews, whose bibliographies, when taken collectively, form an extensive reference list.

The Klarman Review

In an early review, Klarman (1963) used data collected between 1950 and 1961 to assess the effects of prepaid group practice on hospital use (the term HMO had not yet been coined). A simple comparison of hospitalization days per 1,000 members per year showed that during 1950 the rate of use in prepaid plans ranged between 490 and 685 days, while in Blue Cross plans the rate was 888 days, and for the United States population as a whole the rate was 1,165 days. By the year 1960–1961 the rates in prepaid group practice plans had risen somewhat, ranging from 544 to 730 days, with Blue Cross rates rising markedly to 1,060 days, and the rate for the United States population rising to 1,265 days. Klarman considered nine theoretical explanations for the differences between hospitalization rates in prepaid group practice and in conventional health insurance plans. The explanations included differences in 1) the ranges of benefits; 2) the availability of ambulatory care services; 3) access to hospital beds; 4) the possibility of skimping on medical care in the prepaid plans (by failing either to diagnose or to treat existing medical conditions); 5) physician reimbursement; 6) physician control over hospitalization; 7) the role and use of specialists; 8) the willingness of primary care physicians to provide private home health care; and 9) the length of patient stays. After considering these explanations in turn, he concluded:

Of increasing prominence today is the presence or absence of controls. Controls take various forms and may be carried out by

salaried physicians, by subscribers confronted with financial deterrents, or by self-insured plans in which the members actively cooperate; or controls may, in effect, be imposed by lack or inaccessibility of hospital beds. The organizational framework of group practice may constitute a source of control over hospital use, as well as a vehicle for providing ambulatory services. (Klarman, 1963:963-964)

In essence, Klarman attributed the effect to the differences in the controls that influence the organization, the physician, and the patient.

The Weinerman Review

Reviewing basically the same data as Klarman, Weinerman (1964:880) focused on patients' perceptions of group medical care, arguing that "the proof of group practice must lie, after all, in the satisfaction of those who use its special type of service." He found that those enrolled in prepaid group practice were significantly more apt to express complaints concerning "waiting times, inadequate explanations by doctors, difficulty in getting house calls, and lack of interest in the patient as a person" (Weinerman, 1964:885). From these data he distilled the nature of the patient perception problem and cast it in sociological terms: "The local practitioner is pictured as more ready to accommodate the patients' wants . . . whereas the structure of group practice is seen as bending the previously conditioned member to its doctor-oriented rules of procedure" (Weinerman, 1964:886). Weinerman (1964:887-888) went on to sketch the implications of the problem: "Most importantly, the attitudes of patients—whether rational or not—profoundly affect the degree to which the program succeeds in its function, that of protecting its members' health. . . . The uneaten specialty on the dinner plate in the most excellent of restaurants makes little contribution either to nutrition or appetite."

In essence, Weinerman was offering a radically different explanation for the smaller number of hospitalization days in prepaid group practice. He implied that it may not be the greater efficiency in prepaid group practice, but the fact that such alternative health care delivery systems were just so much less "acceptable" (i.e., in terms of the expected and traditionally personal patient-practitioner relationship) that they resulted in a decline in health service utilization.

The Donabedian Review

Donabedian's (1969) review is more comprehensive than either Klarman's or Weirnerman's, and was prepared without reference to them in order to maintain an independent view. Donabedian studied choice of plan, subscriber satisfaction, utilization of ambulatory and hospital services, and the quality of care. He concluded that the major reasons that individuals enroll in a prepaid group practice plan (PGP) (when given the choice through employee benefits) are geographic proximity, not having a private physician as a regular source of care, not being ideologically opposed to "socialized medicine," and the wider range of benefits. Once the employees had chosen membership in prepaid group plans, they were apt to complain about the impersonality of care, the clinic or charity medical care atmosphere, long waiting periods to see a physician, and the difficulty of obtaining house calls. On the other hand, the same subscribers felt that they received good quality medical care in the prepaid group practice, but credited the quality to the availability of technical, diagnostic, and consultative resources rather than to the quality of the physicians themselves. With regard to utilization, Donabedian (1969:11) argued that "the key question is not what is the level of utilization that is associated with any system of organizing care but what precisely happens to utilization and why." In other words, the utilization question should focus on whether the utilization rate is appropriate rather than on what the actual rate is. After an analysis stratified according to disease category (even though the data did not provide a clear answer to the question), Donabedian (1969:15) concluded that "the findings are consistent with the conclusion that [conventional insurance plans] overhospitalize for the common respiratory conditions and the more minor surgical conditions such as benign neoplasms, tonsillectomies and accidental injuries."

In addition to showing that the prepaid plans reduced costs (through lowered hospitalization rates) while maintaining appropriate levels of utilization, the data indicated further cost reductions from the substitution of cheaper ancillary services for the more expensive physician services (also where appropriate). Overall, Donabedian (1969:24-25) presented a very strong case for "the capability of prepaid group practices to achieve a more rational pattern in the use of medical resources, its ability to control costs, and the greater protec-

tion it generally offers against the unpredictable financial ravages of illness."

The Greenlick Review

Building on the earlier reviews, and making ample use of the massive data resources of the Kaiser Foundation Hospitals' Health Services Research Center, Greenlick (1972) assessed the impact of prepaid group practice on American medicine. As in the previous studies, he concluded that comprehensive coverage at a reasonable premium was the major attraction bringing subscribers into prepaid group practices. Greenlick (1972:110) also found that "the expenditures for providing medical care services for a total population covered by prepaid group practice programs are less than the expenditures for care to similar populations covered in the traditional individual fee-for-service system." The reduced cost, however, could not be attributed to efficiencies of scale, but rather arose from "system efficiencies," such that "by integrating the financing and the organization of medical care, PGP can reduce incentives for the physician or the population to prefer that equivalent services be provided on an in-patient rather than on an out-patient basis." In essence, Greenlick concluded that although all the data were not yet in, it was clear that prepaid group practices had several major advantages over conventional plans, the advantages stemming from the different incentives placed before physicians and patients alike.

The Roemer-Shonick Review

Building on Donabedian's, Klarman's, and Weinerman's reviews, and updating them with data from more recent studies, Roemer and Shonick (1973) prepared an extensive review of HMO performance. They focused their review on subscriber composition, participation of physicians, utilization rates, quality assessments, costs and productivity, health status outcomes, and patient attitudes. After reviewing the data, they concluded that

the "prepaid group practice" (PGP) model of HMO continues to yield lower hospital use, relatively more ambulatory and preventive

service, and lower overall costs (counting both premiums and out-of-pocket expenditures) than conventional open-market fee-for-service systems. Economies of scale are still not proved. New data point to reduced disability from the PGP model of HMO, as well as to more favorable consumer attitudes (based mainly on the economic advantage, in spite of certain impersonalities of clinics) than exist toward conventionally insured private solo practice. The medical care foundation [individual practice association] . . . has yielded some evidence of economies in physician's care, but none in hospital use. (Greenlick, 1972:271)

Having built up HMOs as the answer to the crisis in health care, Roemer and Shonick go on to identify the two "principal hazards" that are inherent to the HMO concept: 1) the notion of inequitable "risk" selection, in which the HMO accepts as enrollees only the healthy, for whom the provision of health care is relatively inexpensive; and 2) the provision of poorer quality care through skimping. Moreover, these authors warn that once HMOs move into the mainstream of American health care it will become even more important to maintain vigilance to detect these two hazards, because both the critical approach to a new idea and the self-regulation of a new industry will wear off.

The Gaus et al. Review

Gaus et al. (1976) examined enrollment selectivity, utilization of services, accessibility of care, and patient satisfaction in ten HMOs serving Medicaid patients as opposed to the fee-for-service Medicaid population. With respect to utilization, Gaus et al. found a significant difference in hospitalization only for staff and group models, not for individual practice associations (IPAs). Gaus et al. (1976:3) concluded "that capitation payment to an HMO alone is not significant enough to produce major changes in utilization and that organized multispecialty group-practice arrangements with largely salaried physicians may be more significant." This indicates that some incentives (those related to the organizational delivery of care) are more effective than others (those related to financing). Gaus et al. also found that original health status, use of ambulatory services (including preventive services), patient satisfaction, and access were remarkably similar in both the general Medicaid population and the HMOs.

The Luft Reviews

In three recent papers, Luft (1978a, 1978b, 1979, 1980a; see also his 1979 paper and his forthcoming book, 1980b) has presented the most detailed and critical reviews of the literature on HMO performance to date (even though he makes the serious mistake of assuming that the results of each study should be given equal weight in assessing HMO performance). In particular, from a comprehensive review of the primary literature he has focused on the performance issues of how HMOs actually save money, why they appear to provide more preventive services, and whether they lower the rate of growth of medical care costs. Demonstrating that total costs (both premiums and out-of-pocket expenditures) are from 10 to 40 percent lower in HMOs than in conventional health insurance plans, Luft (1978a:1336) notes that "most of the cost differences are attributable to hospitalization rates about 30 percent lower than those of conventionally insured populations . . . due almost entirely to lower admission rates. . . . There is no evidence that health maintenance organizations reduce admissions in discretionary or unnecessary categories; rather, the data suggest lower admission rates across the board."

According to Luft, there are three possible interpretations of these data: 1) given that discretionary care exists in all categories of hospitalization, an effective HMO may limit discretionary use across the board; 2) self-selection may have the effect that healthier people, who don't need as much hospitalization, come into HMOs; and 3) HMOs may be skipping at the same time that conventional insurance plans "overtreat" discretionary cases. With regard to the issue of whether or not HMOs provide more health maintenance than conventional insurance plans, Luft (1978b:163-164) concludes that "the greater use of preventive services by HMO enrollees appears to be attributable to their better financial coverage, not the preventive care ideology [purported to exist in HMOs]. When people have full coverage for 'preventive' ambulatory visits, they have at least as many, if not more, services under the F[ee] F[or] S[ervice] system than in an HMO. These results are entirely in accord with data for hospitalization—HMO enrollees seem to get fewer services if everything else is held constant."

Focusing on the ability of HMOs to reduce the growth rate of medical care costs, Luft (1980a:1) concludes that "since the early

1960s, total costs for HMO enrollees have grown at a slightly lower rate than for people with conventional insurance coverage. Hospitalization rates show substantial reductions within specific HMOs over a 20 year period." Some of the HMOs' long-term reductions in hospitalization, however, may merely reflect changes in the age-sex mix of their enrollment populations.

Summary

In general terms, these reviews seem to agree on five points: 1) Hospitalization rates in HMOs are up to 45 percent lower than those in conventional insurance systems (this is clearly the case for PGP models, although it is not so clear for IPA models). 2) Total costs are less in HMOs than in conventional insurance systems (again, largely because of lower hospitalization rates, and more pronounced in PGP than in IPA models). 3) Seemingly higher levels of preventive care utilization in HMOs may actually be a reflection of the more extensive coverage that they offer in comparison with conventional health insurance plans. 4) HMO enrollees tend to be more satisfied with the technical aspects of the medical care they receive than are those in conventional insurance plans, but the conventionally insured are more satisfied with their patient-practitioner relationships. 5) Although the evidence on the quality of care received in HMOs is not complete, the quality appears to be at least equal to, if not better than, that received in the average conventional insurance plans. Nonetheless, the most important and agreed-upon point to emerge from these reviews is that *although reduced costs and lower hospitalization rates in HMOs are rather well documented, we still do not know how they are achieved.*

Defining HMOs and Their Incentive Structures

Although the nine reviews discussed above claim to have accurately assessed the performance of HMOs, there are two general sets of problems that they (and the studies on which they are based) have failed to consider in sufficient detail. That is, in order to accurately assess the performance of HMOs, and especially to determine their potential for a major role in any form of national health insurance (NHI), one must first be cognizant of the various problems that

preclude their simple, direct assessment. Specifically, one must consider the definitional problems of what an HMO actually is, distinguish some of the basic types of HMOs from each other, and delineate the general and specific structural incentives and disincentives that affect HMO performance.

Definitional Problems With the HMO Concept.

Like the confusion and controversy surrounding NHI (see Wolinsky, 1980), much of the debate and furor over HMOs may be traced to definitional problems, as a result of which different groups *talk by* rather than *talk to* each other. This confusion occurs because the historical conception of HMOs as Kaiser Health Plan (or Ross-Loos) groups, the legislative definition of HMOs provided by PL 93-222, and the contemporary definition of HMOs as any prepaid health care delivery system, do not agree. To be sure, the PL 93-222 (1973:2) definition, with the DHEW (1974) modifications in brackets, is the most specific: "For purposes of this title, the term 'health maintenance organization' means a legal entity which (1) provides [or arranges for the provision of] basic and supplemental services to its members in the manner prescribed by subsection (b), and (2) is organized and operated in the manner prescribed by subsection (c)." Subsection (b) contains four subsections and subsection (c) contains eleven, all in bureaucratese, and all followed by an entire section of definitions. Thus, although the PL 93-222 definition (as amended by PL 94-460) is the most specific, it is generally used only to distinguish between federally qualified and nonqualified (for feasibility grants and guaranteed start-up loans) HMOs, because it is so cumbersome.

The traditional Kaiser Health Plan definition of HMOs (cf. Greenlick, 1972), while somewhat easier to comprehend than PL 93-222's definition, is not without its own problems. In the main, this traditional definition suggests that all HMOs are closed-panel, hospital-based, group practice models, with enrollment restricted to members of the founding industrial groups (e.g., the Kaiser shipbuilding industries) and subsequent corporate sponsors. The proliferation of HMOs not sponsored by industry, of open-panel plans, and the enrollment of the unemployed and aged under Medicaid and Medicare contracts make this traditional Kaiser Plan imagery less than optimal.

Accordingly, a more general definition is necessary in order to

include the numerous variations on the HMO theme that are currently in existence. Luft (1978a:1336) has used the following definition, which is also well suited to the present paper: "An organization will be considered an HMO if it assumes a contractual responsibility to provide or assure the delivery of health services to a voluntarily enrolled population that pays a fixed premium that is the HMO's major source of revenue." This definition allows us to focus on the larger issue of the performance of HMOs in general, and also allows us to compare and contrast the performances of the specific types of HMOs.

A Typology of HMOs

As we have already seen, there is a considerable amount of variation within the general category of HMOs. Nonetheless, it has become common for health services researchers, policy makers, and the public to delineate only two grossly distinct types in order to simplify comparisons (cf. Roemer and Shonick, 1973). These two types are generally referred to as "prepaid group practices" (PGPs) and "individual practice associations" (IPAs). In PGPs, all of the physicians are members of the same group practice and are reimbursed on the basis of a fixed salary or salary plus profit-sharing, with the group usually servicing the HMO exclusively, and either owning or contracting with the hospital(s) to which its patients are admitted.

On the other hand, the IPA is a loose federation of independent, individual physicians who agree to treat patients enrolled in a third party's HMO (frequently jointly sponsored by medical societies and insurance companies) in their own private offices, being reimbursed by the HMO on a fee-for-service basis (usually less an overhead discount rate of from 10 to 40 percent), and having less than 10 percent of their total patient load coming from the HMO. In essence, the difference is that, although both types of HMOs provide health care to their patients in a prepaid fashion, physicians in PGPs typically are salaried while physicians in IPAs are reimbursed on a fee-for-service basis, and health care is generally delivered from one central location (or satellites thereof) in PGPs, but is delivered out of the individual offices of the independent physicians in IPAs.

Such an arbitrary dichotomization of HMOs into PGPs and IPAs, however, is quite misleading in that there may well be more variation

within each type than between types. Hester (1979:406) has argued that such an arbitrarily dichotomous conceptual model of HMOs is "greatly oversimplified. It neglects key characteristics of the internal structure of those institutions and uses aggregate measures of both inputs and outputs that often blur essential differences in performance." Empirical support for Hester's statement may be found in an analysis of the data reported in the *National HMO Census of Prepaid Plans, 1978* (Department of Health, Education, and Welfare, 1979). Regression analyses of those data reveal that, on the average, group model PGPs (essentially a group practice in which the physicians have a proprietary interest) experience 57.3 more days of hospitalization per 1,000 members than do staff model PGPs (where physicians' services are contracted on a straight salary basis, and where they have no proprietary interest). This difference is both statistically and substantively significant.

Even more compelling reasons than these differences in health outcomes are the theoretically significant structural input differences among HMOs. Table 1 briefly highlights this point by identifying only eight different HMO types according to their different structural configurations. As Table 1 indicates, not all HMOs are alike nor is it simply a case of PGPs versus IPAs. Although this point may seem obvious, it has been seriously overlooked in each of the most often cited reviews of HMO performance (cf. Luft, 1980a, 1978a, 1978b, 1979, 1980b; Gaus et al., 1976; Roemer and Shonick, 1973; Greenlick, 1972; Donabedian, 1969; Weinerman, 1964; Klarman, 1963). As a result, those reviews inevitably suffer from a considerable amount of "conceptual measurement error," because they (and the studies on which they are based) do not isolate the individual effects of the different structural incentives and disincentives of each HMO on its own performance. Accordingly, the five points on which the nine reviews do agree are rather difficult to accept, let alone interpret.

Structural Incentives and Disincentives for HMO Performance

The tandem goals of HMOs are to provide comprehensive health care of high quality to the members, and to provide that care as efficiently as possible. In other words, the *ideas* behind HMOs are *comprehensive health care* and *cost containment*. The principal design characteristic that

TABLE 1
Eight Different Types of HMOs

Structural Characteristics	HMO Models*							
	1	2	3	4	5	6	7	8
HMO owns the hospital, which serves only HMO subscribers	+	+	-	-	-	-	-	-
Physicians are:								
a) Salaried staff of HMO	+	-	-	-	-	-	-	-
b) Salaried group practice of HMO	-	-	+	-	-	-	-	-
c) Group practice legally separate from the HMO	-	+	-	+	-	-	-	-
d) One established group practice contracting their services to the HMO while maintaining their private group practice	-	-	-	-	+	-	-	-
e) Several established groups independently contracting their services to the HMO while maintaining their private group practices	-	-	-	-	-	+	-	-
f) Fee-for-service practitioners who contract individually to serve HMO subscribers while maintaining their private practice for other patients	-	-	-	-	-	-	+	-
g) Fee-for-service practitioners who contract through a separate legal entity to serve HMO subscribers while maintaining their private practices for other patients	-	-	-	-	-	-	-	+

Source: Adapted from "A Report from the Pennsylvania Medical Care Foundation, Prepaid Health Care Primer for Practicing Physicians: Health Maintenance Organizations and Individual Practice Associations," 1978. *Pennsylvania Medicine* 81 (9):10-19.

*Examples of the eight models are: 1) Group Health Plan of Puget Sound (a classic "staff" model). 2) Kaiser Foundation Health Plan (a classic "group" model). 3) Harvard Community Health Plan. 4) Genesee Valley Group Health Association. 5) Med Center Health Plan. 6) Philadelphia Health Plan (a classic "network" model). 7) Employers Mutual of Wausau. 8) Forbes Health Maintenance Plan of Pittsburgh (a classic "IPA" model).

makes HMOs more likely than conventional health care delivery systems to deliver comprehensive health care is the fact that, in the HMO, health care is provided through virtually unlimited access to a complete health care system, with the consumer's cost for that unlimited access being fixed and prepaid. Similarly, the principal design characteristic that makes HMOs more likely to achieve cost containment is that the HMO is placed on a fixed budget, out of which it must meet all the expenses associated with providing the comprehensive health care. Underlying these two design characteristics is the premise that both goals of the HMOs (comprehensive health care and cost containment) can best be achieved by using a systems perspective in which all of the inputs are responsive to a centralized administration. In essence, in order to provide cost containment and comprehensive care, each health care delivery system (i.e., each HMO) must be well integrated in the structural sense. This is the case for all HMOs at the general level, but the extent of structural integration (i.e., structural incentives and disincentives that influence the delivery and consumption of health services) differs markedly from HMO to HMO.

For example, in HMO model 1 (from Table 1) the HMO assumes a supervisory and administrative role with the hospital and the physicians. As a part of this process various incentives (and disincentives) are brought into play, which ultimately affect the delivery and consumption of health services through their promotional or prohibitive nature. For physicians, these incentives may be external or internal. External incentives include a mandatory case review by colleagues to verify the need for a hospitalization recommended by the physician handling the case. Internal incentives include tying the physician's income to the HMO's profit and loss statement (e.g., when physicians are placed at risk for the fiscal health of the plan, after receiving a fixed basic salary). Under such conditions physicians may be more hesitant to hospitalize, especially in proprietary situations where skimping may be financially rewarding. As one physician remarked (see Enright, 1979:127), "Doctors are accustomed to hospitalizing people. When they're put at risk, they begin to examine what they're doing wrong."

A somewhat different picture illustrates HMO model 8 (from Table 1). In this IPA model the HMO does not assume a strict supervisory and administrative role over the hospital. In fact, it merely contracts with the hospital for the right to bring its patients there. Similarly, there are only very weak incentives for physicians to exercise cost

containment, as they are not at risk for the financial success of the hospital, nor are their prospective hospital admissions so subject to their colleague's scrutiny. Accordingly, the IPA is not likely to be as successful at cost containment as is the PGP and, within categories of IPAs and PGPs, there will be a considerable variation in performance (e.g., hospitalization rates) because of different incentives and disincentives.

Specific Incentives and Disincentives

The general realm of incentives and disincentives operative in an HMO includes three categories based on who (or what) actually receives the impact of the incentive: 1) organizational incentives, 2) physician incentives, and 3) patient incentives. In theory, the principal *organizational incentive* is the capitation system of reimbursement. That is, for each of its members the HMO receives a fixed premium in return for providing all the health care the member may need. Accordingly, the HMO must stay within the budgetary constraints imposed by the capitation system, because it may not levy any additional charges. In practice, however, the principal organizational incentive is whether or not the HMO is at risk for hospitalization, and the extent to which it prepares for such risk. Although by definition all HMOs are at risk for hospitalization (as well as for all other expenses incurred while caring for their members under the capitation system), the true extent of an HMO's risk, and the manner in which it employs incentives to reduce hospitalization (and thus reduce costs), vary considerably. For example, the HMO may stress the increased use of preventive ambulatory or outpatient care in order to decrease hospitalization rates. The HMO may also try to reduce discretionary hospitalization by having all potential hospitalizations recommended by any individual physician reviewed for merit by a panel of other physicians. Finally, the HMO may reduce hospitalization by restricting the supply of hospital beds. The expected net effects of all of these organizational incentives is to reduce costly hospitalization while increasing relatively inexpensive preventive ambulatory or outpatient care.

There are two major types of *physician incentives* in HMOs: 1) the effect of their being placed at risk for the delivery of care; and 2) the effect of their being reimbursed on a salary system rather than on a fee-for-service basis. Being placed at risk for the provision of care

essentially places physicians in a proprietary or profit-sharing situation. Under profit-sharing it is to the physician's advantage to decrease expenses, which can be most easily done by decreasing that very expensive item, hospitalization, especially since most hospitalization is initiated by the physician (cf. Andersen and Anderson, 1979; Fuchs, 1974; Wolinsky, 1980). Like the HMO organization itself, physicians have three major ways to reduce hospitalization. 1) They may increase the use of preventive care (reducing the need for hospitalization by early detection and treatment). 2) They may substitute outpatient care and procedures for inpatient care. Or 3) they may choose to reduce the discretionary use of hospital services, such as expensive but unnecessary laboratory tests, or to reduce discretionary hospitalization itself, for such operations as tonsillectomies, hysterectomies, appendectomies, and cholecystectomies (LoGerfo et al., 1979). Any of these approaches should have the net effect of reducing costs. The other major physician-oriented incentive (actually a disincentive) occurs when physicians are reimbursed through a salary (or salary plus profit-sharing) system rather than on a fee-for-service basis. Under the salary system (which is the physician's analogue to the organizational incentives of the capitation system), the physician receives the same income regardless of the number of times he or she sees a given patient, either in the office or in the hospital. As a result, under the salary system there should be a decrease in utilization of hospital and ambulatory care, since any discretionary service utilization represents a diseconomy for the physician. Accordingly, under a salary system the use of hospital and physician services should be less than in a fee-for-service system, all other things being equal.

The major *patient incentive* is the elimination of the out-of-pocket costs (either completely, or retaining only small coinsurance payments such as \$2.00 per visit) usually associated with the consumption of health services. As a result of eliminating the financial barrier of out-of-pocket costs, patients tend to use more preventive and outpatient care (at least during their first association with the HMO), as well as to increase their demand for discretionary services. This should result in a considerable increase in overall ambulatory care, and a net long-term decrease in hospitalization (which, for those who had been conventionally insured, may have had relatively low out-of-pocket costs associated with it even before the HMO experience). Moreover, by eliminating out-of-pocket costs the potential for moral hazard is

greater in HMOs than in conventional insurance plans. In other words, when the financial barrier of out-of-pocket expenses for physician utilization is removed, utilization of physicians' services should increase, all other things being equal.

As indicated above, we can directly assess the effect of any specific incentive only if the effects of the other incentives are held constant (or controlled). Unfortunately, it is difficult to study HMOs with the randomized experimental designs necessary to accomplish this task—the Seattle Prepaid Health Care Project (Diehr et al., 1976) notwithstanding. As a result, there have been few research situations (and few, if any, HMO performance reviews) in which clear and unambiguous tests of these incentives have been made. Therefore, future HMO performance studies and reviews will need to identify the different structural incentives operating, as well as the extent of their operations. Otherwise, combinations of countervailing incentives will continue to go undetected, obscuring the analysis and resulting in artificial evaluations of HMO performance. The results of several recent studies of health service utilization support this need to identify the structural characteristics of the specific health care delivery system, when health outcomes are being examined (Dutton, 1978, 1979; Kronenfeld, 1978; Shortell et al., 1977; Williams et al., 1978; Wolinsky, 1978).

Methodological Problems in Evaluating HMO Performance

In addition to the general set of problems in defining HMOs and identifying their incentive structures, there is a second general set of problems of a more methodological nature, which the nine HMO performance reviews (and the studies on which they are based) have also failed to consider in sufficient detail. Specifically, there are three methodological issues that warrant further attention before any analytical review of the literature: 1) the effects of adverse self-selection and other population differences; 2) the development and selection of the health outcome measures used to assess HMO performance; and 3) the efficacy of survey versus plan-audit data collection techniques. Until these issues are carefully considered, summary statements con-

cerning HMO performance will remain ambiguous, if not downright misleading.

Adverse Self-Selection and Other Population Differences

The first methodological issue to be considered is the two-sided question of adverse self-selection and other population differences. That is, if it is not possible to arbitrarily and randomly assign subjects to HMOs that represent the different combinations of incentives and disincentives, or to traditional health care delivery systems, the comparability of subjects in the different systems must be established. The possibility that voluntary HMO enrollees were significantly different from those not electing (volunteering) to enroll in HMOs was first suggested by Bashshur and Metzner (1970:106), and has come to be known as *the risk-vulnerability hypothesis*:

Persons in the younger age groups, single persons and those with no dependents and those of lower socioeconomic status . . . may be viewed as having a lesser economic "stake" and therefore feeling themselves less vulnerable to serious economic loss in meeting health needs. Older persons, those with dependents, on the other hand, may be considered to run a greater economic risk or to be greater health risks (or both) and hence to be more vulnerable.

According to this hypothesis, the more vulnerable, either in the economic or the health sense, are more likely to enroll in HMOs than are the less vulnerable. If this is the case, HMOs might appear to be less cost efficient merely because they commence with a more morbid or potentially morbid population requiring more care (Berki et al., 1977a, 1977b, 1978; Ashcraft et al., 1978). For an excellent review of the relevant literature, see Berki and Ashcraft (1980).

The recent data on the effects of voluntary disenrollment from HMOs make the risk-vulnerability hypothesis even more important. Wollstadt et al. (1978) have shown that individuals without dependents voluntarily disenrolled from HMOs at a rate twice as fast as those with families. Moreover, among those with families, the larger the family size (i.e., the more economically vulnerable), the less likely was voluntary disenrollment. Those who voluntarily disenrolled also

had a much higher rate of out-of-plan use (indicative of established patient-practitioner relationships) than those who remained in the HMO. Taken together, these data suggest that the voluntary disenrollment process may well be reinforcing the alleged effects of the risk-vulnerability hypothesis in creating noncomparable populations. That is, there may be not only an adverse self-selection process, but also a complementary adverse voluntary disenrollment process.

The other side of this problem is that even if perceived health risk and economic vulnerability are similar throughout the HMOs and the conventional comparison groups, measures of outcome (such as indicators of health service utilization) are not directly comparable unless they are first adjusted for other population characteristics (most notably age, sex, race, and socioeconomic status). Such adjustments are necessary because of the effects these predisposing and enabling factors have on the use of services (Andersen, 1968; Andersen and Newman, 1973; Aday et al., 1980). Thus, without adjustment for the differential distribution of these characteristics across the health care delivery systems to be compared, the effects of the HMO's incentives can not be accurately assessed, and may appear to be contradictory from study to study. This is evident in Table 2, which contains a summary of the unadjusted results of eighteen previous studies of HMO performance, made by calculating the ratio of utilization in the HMO to the utilization of the fee-for-service comparison system (Diehr et al., 1976). Ratios less than unity indicate better HMO performance, and ratios greater than unity indicate poorer HMO performance (relative to the fee-for-service comparison group). Considering the fact that "in 6 of the studies (4, 7, 9, 10, 11, and 12) sociodemographic correlates are either not provided or would suggest lower utilization in general by prepaid group practice patients" (Diehr et al., 1976:1), the ratios in Table 2 reflect how volatile unadjusted and otherwise noncomparable results may be.

Health Outcomes: The Comparison Measures

As we have already seen, performance studies of HMOs have focused, in the main, on hospital and physician utilization (Table 2). In fact, we limited our own explication of the incentives and disincentives that make up the various HMO types to a discussion of what the hypothesized effects of the various incentives would be on physician

and hospital utilization. Such traditionally limited approaches, however, pose two rather serious limitations for any subsequent analysis. *First*, by using only hospital and physician utilization as indicators of health outcomes, *a considerable part of the important conceptual domain of health services is omitted*, including nonphysician ambulatory visits, such as those by registered nurses, nurse practitioners, and physician assistants (which are often substituted for physician visits in HMOs), accessibility of care (time and distance to the physician's office as well as waiting time, both to schedule appointments and on arrival for scheduled appointments), patient satisfaction, the quality and continuity of care, the efficiency of care (economic issues), and the effectiveness of care (how well the delivery system is structurally integrated).

Second, by using only the gross summary measures of physician and hospital utilization, *a considerable amount of measurement error is introduced*. That is, when the only measure of ambulatory care is the gross per capita number of physician visits, or when the only measure of hospital care is either the gross admissions per 1,000 members or the per capita days of hospitalization, it becomes very difficult (if not impossible) to interpret (partition) the results accurately and ultimately evaluate the performance of HMOs. For example, knowing the gross per capita number of physician visits per year does not allow one to determine the effects of the various incentives, because although preventive and outpatient care may increase, the number of physician visits to hospitalized patients and the number of hospitalized patients themselves may decrease in a countervailing fashion. As a result, it would appear that physician visits were unaffected by the incentives operating in the HMO, when in fact the *mix* of physician visits had actually changed markedly. A similar example is provided by hospitalization rates. If the organizational and physician incentives operating in HMOs have the desired effects, those effects should be most prominent in terms of reduced surgical hospitalizations and surgery-induced hospitalization days. Using a gross hospitalization measure, however, might well mask the dramatic nature of the drop in surgical hospitalization. Accordingly, when health outcomes among different HMOs or conventional health care delivery systems are being compared and contrasted, it is necessary to determine the comparability and the degree of measurement error of the health outcome measures under examination.

TABLE 2
 Ratios of Utilization and Cost Measures in Prepaid Groups to Those in Fee-for-Service Independent Practices,
 in Eighteen Studies*

Study	Hospital		Physician		Total Costs
	Admissions	Days	Visits	Surg. Proc.	
1. Densen et al. (1958)	.86	.91	—	.82	—
2. Anderson and Sheatsley (1959)	.57	.47	.92	.65	.90
3. Densen et al. (1960)	.80	.78	—	.79	—
4. Falk and Senturia (1960)	.85	.55	.77	.48	—
	.77	.40	1.61	.52	—
5. Densen et al. (1962)	1.01	1.00	—	—	—
6. Williams et al. (1964)	1.04	1.05	1.32	1.05	.86
	1.11	1.00	1.13	1.61	.89
7. Dozier et al. (1964)	.65	.81	1.06	—	.77
8. Shapiro et al. (1967)	—	.87	.98	—	—
9. Chamberlain (1967)	.60	.53	—	—	—
10. National Advisory Com. (1967)	.68	.69	.94	—	.54
11. Perrott and Chase (1968)	.50	.47	—	.42	—
12. Perrott (1971)	.49	.49	—	.45	—
13. Robertson (1972)	.49	.40	2.13	.33	—
14. Hastings et al. (1973)	.80	.76	1.48	—	—
15. Hetherington et al. (1975)	1.05	.42	1.07	—	.89
16. Broida (1975)	1.54	1.41	2.29	—	—
17. Reidel et al. (1975)	.57	.58	—	—	—
18. Wersinger et al. (1976)	.64	.63	—	—	—

*Based on material adapted by Thomas Dolan from an earlier compilation by Goldman and Steinwald (1973), and published in Diehr et al. (1976). Areas of Study in Table 2.

1. Health Insurance Plan of Greater New York (HIP) and Blue Shield (BS) of New York City.
2. HIP and Group Health Insurance, Inc. (GHI), of New York City.
3. HIP and GHI of New York City.
4. Steelworkers enrolled in three health care plans: Kaiser-Permanente (KP) Medical Care Program in California, Blue Cross (BC) plans in various locations, and commercial carrier plans in various locations.
5. HIP and union-administered fee-for-service plan in New York City.
6. KP-San Francisco Bay area, BC/BS New Jersey, and General Electric health plan in Utica, New York. A possible explanation for the comparable hospital utilization in the three plans is that the sample size was inadequate for reliable comparisons of hospital utilization. In addition, there were large geographical differences nationally among the three patient populations.
7. KP-California and BC/BS California. Patient discharge rate is shown in place of hospital admissions rate.
8. Old Age Assistance recipients in New York City served by HIP and fee-for-service practitioners.
9. Various unidentified prepaid group-practice plans nationwide, BC subscribers nationwide.
10. KP-California and population at large, California. Physician visits data compare KP-California with residents of the western U.S.
11. Federal employees enrolled in various prepaid plans and BC/BS plans nationwide.
12. Federal employees enrolled in Group Health Association (GHA) and in BC plan, Washington, D.C. Surgical procedures data compare federal employees in prepaid group practice plans and in BC plans nationwide.
13. Public school teachers enrolled in an unidentified prepaid group practice plan and an unidentified nonprofit plan.
14. Employees of Algoma Steel Corporation, Sault Ste. Marie, Ontario, enrolled in either the Sault Ste. Marie and District Group Health Association or an indemnification plan administered by the Prudential Insurance Company of America.
15. Subscribers to two prepaid group practice plans, two commercial carrier plans, and two "provider" plans (including BC/BS), in California. The authors attribute the relatively low admission rate of the commercial plans to their acceptance of fewer "high-risk" patients.
16. Prepaid enrollees compared with fee-for-service patients, Marshfield Clinic, Marshfield, Wisconsin.
17. Federal employees in the Washington, D.C., area enrolled under the high-option Blue Cross-Blue Shield Government-Wide Service Benefit Plan compared with those enrolled in Group Health Association, Inc.
18. Enrollees under a foundation-type plan, Health Watch, and a BS/BC-sponsored multispecialty prepaid group practice, Rochester, New York. Rates exclude maternity-related and psychiatric admissions.

*Data Collection Problems:
Patient Surveys versus Plan Audits*

To collect data on health outcomes, either the patient survey or the plan-audits technique is used, each of which has its own methodological problems. In the plan-audits technique, data are collected by abstracting information from the record-keeping system of the health plan. In other words, an archival record search is made, and whatever utilization has been recorded by the plan for a given individual becomes that individual's utilization data for the HMO performance study. The single most important problem with the plan-audit technique is that only in-plan utilization is counted. Out-of-plan utilization (primarily ambulatory care services), which occurs to some extent in all plans and has been estimated to be as high as 53.9 percent of total utilization under extreme circumstances, goes unmeasured (Greenfield et al., 1978; Corbin and Krute, 1975; Bashshur et al., 1967; Freidson, 1961). Therefore, HMOs may appear to be more efficient (or not to be substituting more ambulatory care for less hospital care) because the plan-auditing technique underreports the actual amount of health services used by plan members.

A second important problem with using plan-audit data occurs when data from different plans are to be compared. It is quite likely that each HMO institutes its own data collection and retrieval systems, each system being based on somewhat different definitions and having rather different degrees of completeness and accuracy. Accordingly, any observed differences in plan-audit data between HMOs may well be a function of the different ways in which records were kept or audited, rather than reflections of "true" differences in HMO performance.

On the other hand, the patient survey technique, while eliminating the problem of measuring the out-of-plan use of services (but not necessarily the problem of noncomparability), is much more expensive and time-consuming. In addition, patient surveys have their own special problems, including those in general recall, in the reporting of detailed specific information such as initial diagnosis, determining whether the services were used for preventive or for restorative care, and what specific treatment regimen was prescribed. In other words, while the patient survey technique captures out-of-plan utilization, it introduces general and specific recall problems (which, even though

they occur in the form of intercoder reliability problems in plan-audits, are at least relatively constant across the smaller number of doctors and nurses, compared with the variance across the larger number of enrollees). Therefore, when the performances of HMOs are compared with each other or with those of conventional health care delivery systems (as control groups), the methods of data collection must be comparable.

A Review of HMO Performance

Having defined and identified the incentive structures and the methodological problems involved in studying HMOs, we now come to the analytic review of HMO performance itself. The most comprehensive and direct tack would be to collect the necessary and detailed information for all HMOs currently in operation. Such an approach, however, is clearly beyond the limits of the present paper. A more feasible approach is to briefly describe the current population of HMOs, by means of the most comprehensive census data available, and then to analytically review the recent primary literature on HMO performance in light of the issues described above.

A Brief Survey of Prepaid Group Practices

The most comprehensive and timely source currently available for data on all HMOs is the *National HMO Census of Prepaid Plans, 1978*, compiled by the Office of Health Maintenance Organizations (Department of Health, Education, and Welfare, 1979). The HMO census contains relevant data (albeit somewhat limited and not necessarily uniform) for each of the 203 HMOs in the United States as of November 30, 1978 (since that time the number of HMOs has grown by approximately 15 percent, with most of the newest HMOs being IPAs). These data indicate that total membership in HMOs has risen to 7,470,963 individuals, of which 1) more than two-thirds are in the larger HMOs (those with 100,000 or more members), 2) nearly three-fourths are in the older HMOs (those in operation for ten or more years), and 3) nearly half are in California. On the average, the

monthly premium for families in all plans was \$95.32 (covering the most comprehensive, high-option plan available).

In addition to these more general descriptive statements, Table 3 contains the means, medians, and standard deviations for the general membership, Medicare membership, Medicaid membership, federal employees membership, number of years in operation, hospitalization days per 1,000 members per year, and number of physician encounters per member per year by HMO type (loosely categorized as staff, group, or IPA models). The data in Table 3 show that these general characteristics of HMOs are as varied as the incentives that comprise their organizational structures. Moreover, comparing the relative sizes of the means, medians, and standard deviations clearly indicates that these distributions are quite skewed, providing more support for the argument that not all HMOs are alike and that they may not be directly compared without using the appropriate statistical controls.

Multivariate analyses of these data reveal that when HMO performance is defined as hospitalization days per 1,000 members per year, a significant prediction model can be obtained. That is, using dummy variables for group models, IPA models, and federal qualification as an HMO yields a regression equation having an explained variance of .11, in which the overall and individual effects of the dummy variables are all significant at the .05 level:

$$\begin{array}{l} \text{Hospitalization days} \\ \text{per 1,000 members} \\ \text{per year} \end{array} = 425.6 + 57.3 \text{ group} + 78.7 \text{ IPA} - \\ 50.4 \text{ federal qualification} + \text{error.}$$

In other words, even though the HMO census data are far less than optimal, (i.e., the data are not adjusted for different age and sex distributions, nor is the classification of HMO types an exceedingly rigorous one), they indicate that HMO performance ranges from 375.2 days per 1,000 members per year in federally qualified staff models (where the greatest combination of incentives is in operation), to 504.3 days per 1,000 members per year in nonfederally qualified IPA models (where the least amount of incentives is in operation), with qualified and nonqualified group models in between at 432.5 and 482.9 days, respectively, per 1,000 members per year. These data provide additional support for the need to classify HMOs in terms of their combinations of incentives.

An Analytic Review of the Recent HMO Performance Literature

We now proceed with the analytic review of the recent literature (1977 or later) on HMO performance, in light of the issues raised above. We focus on the recent literature for three reasons: 1) given the large volume of the extant literature, some limitations are necessary for practical reasons; 2) previous reviews have summarized the literature up to the last few years; and 3) the increased amount of information commonly provided in the more recent HMO performance studies permits a more thorough review of how certain conclusions have been reached, why they may be in error, and how they can be adjusted. To maximize continuity, these studies are reviewed in loosely defined groups based on the research project or institutional affiliation of the authors.

The Stanford Project. In an attempt to assess the effects of supply-side and demand-side incentives on the use of physician services, Scitovsky et al. (1979) conducted a twelve-month study of 4,200 individuals enrolled in two prepaid plans. All respondents were employees (or their dependents) of Stanford University, and were faced with a triple-choice situation (with a conventional Blue Cross-Blue Shield plan as the alternative). One prepaid plan was a Kaiser plan (a closed-panel, group model with completely integrated facilities), and the other was a fee-for-service group practice that was not at risk for hospital costs, with only 15 percent of its revenue coming from prepaid plans. The Kaiser plan had a token copayment for physician office visits (\$1.00); the other prepaid plan had a 25 percent coinsurance charge for all physician visits. On the one hand, Scitovsky et al. hypothesized that the at-risk physician incentives in the Kaiser plan would result in lower physician utilization. On the other hand, it was expected that physician utilization would be lower in the fee-for-service group practice, because of the 25 percent patient coinsurance disincentive. Unfortunately, the design of the study does not allow these individual effects to be accurately partitioned, so that it is impossible to determine why whatever happened, happened. Nonetheless, Scitovsky et al. report that after allowance for age composition, socioeconomic status, health status, attitudes toward seeking care, length of membership in plan, family size, and plan satisfaction, the mean number of physician visits per person per year is about the

TABLE 3
Means, Medians, and Standard Deviations of HMO Characteristics by HMO Type in 1978*

HMO Types	Characteristic	Mean	Median	Standard Deviation
All HMOs (N = 203)	Enrollment:			
	General	37,923	9,061	159,759
	Medicare (N = 61)	6,170	497	16,743
	Medicaid (N = 35)	6,570	2,792	11,057
	Federal employee (N = 55)	10,608	1,188	28,529
	Number of years in operation	6.566	4.570	8.568
	Hospitalization days per 1,000 members per year	454.535	438.000	134.583
Group Model HMOs (N = 78)	Physician visits per member per year	3.282	3.287	0.951
	Enrollment:			
	General	69,963	3,463	248,192
	Medicare (N = 27)	10,678	1,035	23,824
	Medicaid (N = 9)	13,515	6,504	18,486
	Federal employee (N = 24)	18,336	2,754	39,753
	Number of years in operation	8.274	5.350	8.846
Hospitalization days per 1,000 members per year	462.00	458.000	142.814	
Physician visits per member per year	3.202	3.250	0.9016	

Staff Model HMOs (N = 52)	Enrollment:			
	General	19,186	9,851	37,716
	Medicare (N = 19)	2,033	233	4,237
	Medicaid (N = 14)	3,038	1,693	3,625
	Federal employee (N = 18)	6,736	1,157	16,770
	Number of years in operation	6,450	4,317	8,562
	Hospitalization days per 1,000 members per year	410.710	394.500	112.281
	Physician visits per member per year	3.319	3.350	0.827
IPA Model HMOs (N = 70)	Enrollment:			
	General	15,686	8,000	29,612
	Medicare (N = 15)	3,299	185	7,278
	Medicaid (N = 12)	5,484	2,107	7,561
	Federal employee (N = 13)	1,703	922	2,712
	Number of years in operation	4,919	3,550	8,122
	Hospitalization days per 1,000 members per year	485.476	480.500	135.885
	Physician visits per member per year	3.343	3.367	1.117

*Adapted from Department of Health, Education, and Welfare (1979).

same. After the data are adjusted for physician affiliation (having a plan physician as the regular source of care), however, Kaiser patients use .48 more physician and .79 more physician-plus-paramedic visits per person per year.

On the surface these data may seem to indicate that the Kaiser plan is less efficient, but there are at least three other plausible and competing interpretations that can not be rejected. First, the increase in the use of physician services may reflect a more aggressive substitution of ambulatory care for hospital care in the structurally well-integrated (plan-owned hospital facilities) Kaiser plan. Because hospitalization rates are not reported, such trade-offs can not be documented. Second, the relative effect of the coinsurance patient disincentive of the other prepaid plan may exceed the effect of the well-integrated Kaiser plan incentives, a possibility that can not be assessed with these data, either. Third, the differences in physician use may be a function of differential plan selection (cf. Wolinsky, 1976) based on the uncontrolled effects of social class characteristics (such as occupational prestige) as opposed to the controlled effects of socioeconomic status (such as income). That is, members of the Kaiser plan are predominantly support staff at Stanford, whereas members of the other plan are predominantly faculty. Therefore the different levels of the use of physician services may be a function of different lifestyles, which would not necessarily be reflected by income distributions. In an earlier paper assessing differences among those enrolling in the two plans, Scitovsky et al. (1978) simply declared that this was not the case, and proceeded to eliminate occupation and education from the analysis, concluding that family income and proximity to the health plan's medical center are the major factors affecting choice. Support for a rival lifestyle interpretation, however, comes from the fact that 45 percent more of the fee-for-service plan members (faculty) had a specific plan physician as their regular source of care than did Kaiser members, which may be interpreted as a basic function of their lifestyle. In sum, although the Scitovsky et al. study provides an interesting look at the use of physicians' services in two prepaid plans, methodological shortcomings prohibit any explicit tests of the structural incentives in operation.

The Seattle Project. As part of the Seattle Prepaid Health Care Project, Williams et al. (1979) examined the use of mental health services in a prepaid group practice and in an individual practice

association (in which the physicians were not at risk). Although there were no out-of-pocket costs to enrollees in either plan, the IPA emphasized individual psychotherapy, and the PGP had a more diverse orientation, employing a variety of practitioner and therapeutic modalities. Although enrollees were not randomly assigned, and as a result there were some small (yet statistically significant) differences between the IPA and PGP enrollees, the resulting effects of adverse selection appear to be counterbalancing. Accordingly, these data invite a rather direct comparison. Williams et al. (1979:148) report: "The percent using any mental health services was twice as great in the PGP as in the IPP [IPA], although persons using any services tended to have more visits in the IPP as compared to the PGP. A significantly higher proportion of enrollees were hospitalized in the IPP than in the PGP."

These data support three hypotheses concerning HMOs: 1) In an HMO (especially in the PGP model), paramedical services will be actively substituted (where appropriate) for more costly physician services. 2) The better integrated the hospital services (PGP versus IPA), the lower the number of hospitalization days and admissions. And 3) the better integrated the HMO is overall (PGP versus IPA), the more likely it is to have a health maintenance orientation, as reflected in extending more preventive (e.g., counseling and paramedical mental health) services to its enrollees. Although they do not provide a direct test of Goldensohn and Fink's (1979) hypothesis (that psychiatric treatment reduces the use of other physician's services), the data given by Williams et al. (1979) appear at first to be consistent with it, as do those of Levin and Glasser (1979), in their survey of the use of mental health services.

On closer inspection, however, the data of Williams et al. more clearly support Mechanic's (1979, 1980) hypotheses that increased and internalized psychological or psychosocial bodily concerns result in high levels of symptom sensitivity and reporting, which should ultimately lead to the increased use of physician services. Using other data from the Seattle project, Diehr et al. (1979a:937) demonstrate this phenomenon, concluding that mental health utilizers "consumed more somatic services than other enrollees, even controlling for background variables." Patrick et al. (1978) have also demonstrated that patients with chronic emotional problems, and their families, use more physician services, and the authors question whether reduced

costs can ever be the outcome of such increased emotionally induced physician utilization. Unfortunately, the Patrick et al. study does not provide any cost comparison data with which to evaluate the performance of its PGP.

In another aspect of the Seattle project, LoGerfo et al. (1979) studied the rates of surgical care in the PGP and the IPA. Specifically, they were interested in determining whether differences in the surgical rates existed and, if so, what accounted for them. Calculating exposure-adjusted ratios (IPA:PGP) resulted in gross ratios of 3.8:1 overall, 5.8:1 for tonsillectomies, 6.3:1 for hysterectomies, 1.7:1 for cholecystectomies, and 1.7:1 for appendectomies. These ratios indicate that much more surgery takes place in the IPA than in the PGP. To determine whether or not the differences are due solely to more appropriate discretion in the PGP, LoGerfo et al. (1979:1) used a variety of acceptable (but not definitive) algorithms to isolate only those surgeries that were "necessary, appropriate, or justified." After reviewing the data they found ratios of 2.8:1 overall, 2.8:1 for tonsillectomies, 6.8:1 for hysterectomies, 1.7:1 for cholecystectomies, and 1.4:1 for appendectomies. These ratios indicate that even when the analysis is restricted to justifiable surgery, the rates are significantly higher (for tonsillectomies and hysterectomies) in the IPA than in the PGP. Because the characteristics (including health status) of the enrollees in the IPA and the PGP are very similar, this points to the effect of organizational incentives as the reason for less surgery in the PGP. Because all of these surgical procedures were justified, LoGerfo et al. (1979:12) concluded that "there is substantial evidence to support the contention of underprovision of surgical care in prepaid group practices as an explanation for observed differences in surgical rates."

In short, the PGP appears to have been skimping on services. Although there are some measurement and sampling problems involved in this study (such as the determination of "justifiable" surgery, the rather small number of types of surgical procedures, and the fact that true estimates of the number of people at risk for each procedure are not available), they are not of sufficient magnitude to completely offset the serious implications of skimping. Therefore, these data suggest that although placing physicians at risk for hospitalization costs (as in the PGP, and not in the IPA) eliminates a significant amount of unnecessary surgery, the effects of the incentives are dysfunctional: they go beyond reducing unnecessary services, perhaps far

enough to induce skimping on necessary and justifiable services. Moreover, a subsequent analysis (Diehr et al., 1979b) of the effects of the increased access to medical care (offered by both the IPA and the PGP) on health status produced poorer health status evaluations on four out of five self-report measures (after one year's experience in the HMO). In addition, after one year enrollees were less healthy on all five measures than those in a comparison group (in a limited access, or nonenrolled situation). Unfortunately, the subjective self-report nature of these data severely limits their reliability and validity, let alone their generalizability, as they are highly subjective and susceptible to the low levels of satisfaction with the patient-practitioner relationships typically found in HMOs (relationships that were neither examined nor described).

In a related paper, however, LoGerfo et al. (1978) reached a far more favorable conclusion concerning the quality (process) of care in the PGP—a conclusion consistent with Williamson et al.'s (1979) unpublished summary of the literature on the quality of care in HMOs. Focusing on the care associated with urinary tract infections, LoGerfo et al. (1978:494) found that the quality of care was significantly better in the PGP. Specifically, "there was a greater recording of pertinent history and physical exam items in the prepaid group practice; the prepaid group practice used markedly more urine cultures and slightly more urinalyses in the laboratory evaluation of these patients; and the prepaid group practice physicians tended to use a more uniform and appropriate set of antibiotics."

They also found better quality care processes in the PGP for common infections (LoGerfo et al., 1977) and hypertension management (LoGerfo, 1975). Merging these findings with their earlier study on the appropriateness of hospitalization leads one to conclude that although there may be more of a tendency to skimp (entirely) on the delivery of care in the PGP, once the decision to provide care has been made the quality of the delivered care is better than that of the IPA. An alternative explanation would be that PGPs skimp only on expensive services (i.e., surgery) and not on inexpensive ones (i.e., laboratory examinations). In either case, the implication of skimping remains, although the nature of the skimping (i.e., selective vs. across-the-board) is unclear. Unfortunately, the design of the Seattle project precludes any comparisons of either the PGP or the IPA with a conventionally insured control group (from the patient's perspective).

The Berki Papers. In four papers reporting on different aspects of a large study of a quadruple-choice health plan situation, Berki and his colleagues (Berki et al., 1977a, 1977b, 1978; Ashcraft et al., 1978) have presented the most detailed and sophisticated analyses of a field study to date. Focusing on enrollment choice and the expectation and actual use of services, they employed a variety of statistical controls and analytic approaches, and concluded that

having a private physician as the source of care is the best single predictor, its absence predicting a higher probability of enrollment in the closed, and its presence in the open-panel HMO. Higher risk life state families, younger and with more children, are more likely to join the open-panel plan than the closed or retain BC/BS; higher incomes and larger numbers of chronic conditions appear to have the same effects. Higher levels of health concern, on the other hand, predict a greater probability of choosing the closed-panel plan. (Berki et al., 1978:682)

This indicates the occurrence of statistical interaction such that having or not having a private physician as a regular source of care specifies the effects of health concern and stage of family life cycle on enrollment choice. In the presence of a private family physician, health concern does not affect enrollment choice, whereas family life stage does (the open-panel HMO being chosen). The opposite holds true in the absence of a private family physician (the closed-panel HMO being chosen).

Perhaps more interesting than their actual analysis is the conceptual refinement that Berki et al. offer. Specifically, they divide the risk-vulnerability hypothesis into two parts, financial vulnerability and health risk. Focusing on financial vulnerability (uniquely operationalized as per capita income rather than family income, in order to focus on relative vulnerability), they argue that open-panel HMO selection (i.e., selecting IPAs) versus retention of Blue Cross and Blue Shield (the Blues) is the most clear-cut test, since the difference between these two plans involves only choosing among insurance plans (risk) and not among physicians (providers). Comparisons of per capita income of those retaining the Blues and those choosing the IPA, however, do not reveal significant differences. Comparisons of family incomes do reveal significant differences, al-

though they are in the opposite direction (IPA families have approximately \$2,800 more income and thus less financial vulnerability than Blues families). Unfortunately, Berki et al. ignore their own insights and focus on general comparisons of HMO versus Blues selection (rather than on specific comparisons of IPAs versus Blues), which show that enrollees in any type of HMO have significantly lower per capita incomes than their Blues counterparts. Nonetheless, following their own conceptual breakthrough (comparisons of IPA versus Blues enrollees), they find no support for the financial vulnerability hypothesis. Similarly, Berki et al. (1977b:112) conclude that in terms of illness conditions there is also "no evidence of adverse health risk self-selection in an employed population."

After the enrollees had had one year's experience in the HMO, Berki and his colleagues (Ashcraft et al., 1978) refocused their analysis on satisfaction, utilization, and costs. Following an insightful panel analysis of HMO enrollees who remained in the HMO throughout the study, as well as those who had retained the Blues throughout, they concluded:

Lack of access to and dissatisfaction with previous sources of care distinguished the preenrollment experience of those who selected the closed-panel plans; their postenrollment experience produced increased satisfaction reflecting that their expectations in these areas were met. Continuing enrollees in closed-panel plans were somewhat less satisfied after a year of experience than they were earlier. Those who joined the open-panel did so because of the expanded benefits and financial advantages which, their postenrollment experience showed, were accurately perceived. . . . [C]ontinuing enrollees in both types of plans made fewer illness but more preventive visits; new enrollees used greater numbers of both types of services after enrolling than before. (Ashcraft et al., 1978:14)

These data support three hypotheses concerning HMOs: 1) that the use of preventive care will be increased, which may ultimately result in reduced overall costs (although beyond the diagnosis and treatment of hypertension and diabetes, the data are not very clear on this point); 2) that the overall use of ambulatory services increases (at least during the first year), apparently concomitant with a decrease in hospitalization (although these data are not presented); and 3) that HMO enrollees are more satisfied overall (as a function of increased general

access), but their level of satisfaction declines somewhat over time (probably as a function of less than optimal patient-practitioner relationships in the closed-panel plans).

Kaiser Plan Studies. Focusing on consumer satisfaction, Pope (1978) used multivariate analytic techniques on data from samples of currently active and recently terminated subscribers to the Kaiser-Permanente Medical Care program. He found that, among current subscribers, satisfaction is the highest for those who have a regular plan physician, are older, and live in families who rate their own health as excellent. Pope found that those who terminated their enrollment, as a result of dissatisfaction with Kaiser, are more likely to rate their own families' health as less than excellent, do not have a regular plan physician, and have been local residents for quite some time. Although Pope carefully points out that these results may be a function more of expectations than of the delivery (or nondelivery) of medical care in HMOs (a problem that the data can not resolve), there are two other serious limitations on these data. First, the response rate is rather low among the active subscribers (65 percent) and extremely low (29 percent) among those terminating their membership; only 6 and 16 percent, respectively, did not respond because they could not be reached. Therefore the satisfaction expressed by the terminating sample is likely to be biased, overrepresenting outlier values. Second, because satisfaction data are not presented for a comparable group from the conventionally insured population, the effects of HMOs on satisfaction can not be determined.

Using data on enrollees in Kaiser-Permanente's Oregon region, Freeborn et al. (1977:115) assessed the relative effects of health and of socioeconomic status on the use of ambulatory services. Controlling for the effects of age and sex, they found "health status to correlate more highly than socioeconomic factors with the utilization of services. . . . An exception was the use of preventive services, which was not significantly related to health status measures but rather, for women, to education, and to a lesser extent, income."

Unfortunately, an overreliance on zero-order correlations casts doubt on the results of this attempt to explicate the relative causal effects of health and socioeconomic status on ambulatory care use. The data do, however, suggest that enrollment in an HMO does not bring about equity in the use of preventive services, especially when predisposing characteristics such as education are considered. In ef-

fect, these data indicate that even in the absence of financial barriers, health services (at least preventive ones) will not be used unless there is a concomitant predisposition (or preventive health orientation).

Using data on 3,892 individuals enrolled in the Kaiser Foundation Health Plan of Portland, Lairson and Swint (1978, 1979) focused on the determinants of preventive and nonpreventive utilization of medical care service. Employing a modified version of Aday and Andersen's (1975) framework for the study of access to medical care, Lairson and Swint obtained results from a series of regression analyses (both ordinary least-squares and logistic) indicating that, for nonpreventive services, utilization is a function of health status, age, and sex; for preventive services, utilization is a function of education (positive), income (positive for dependents' utilization only), and age (positive for the young and the old). These data support the interpretation suggested earlier, that although enrollment in HMOs may remove the economic disincentives that serve to restrict the use of preventive services, enrollment in the HMO alone does not produce equity in preventive health behavior. Rather, preventive health service utilization is closely related to lifestyle and the ability to successfully negotiate with the HMO's bureaucratic system, both of which may be proxy-measured by education. In short, once enabling barriers are removed, one has to be predisposed in order to use services. In the preventive services case (when health status is controlled), this reduces to the preventive health orientation, which is directly related to education. Because Lairson and Swint do not control for the effects of different lengths of plan membership, it is impossible to determine whether or not the effect of education on preventive service utilization decreases with continued exposure to the HMO and its alleged preventive health or health maintenance orientation.

Using data collected from three samples of low-income individuals (a group enrolled in the Kaiser Health Plan of Oregon, a matched group of Medicaid recipients, and a group enrolled in the Kaiser plan who also retained their Medicaid eligibility), Johnson and Azevedo (1979b) sought to assess two general hypotheses: 1) that the use of services and the resultant costs of low-income enrollees in a PGP model HMO would be lower than those for a matched Medicaid group; and 2) that the use of services and resultant costs of the low-income PGP enrollees would also be less than those for a matched group of Medicaid recipients who were concurrently enrolled in the

PGP. Their intent was to assess the cost-containment strategy of enrolling Medicaid recipients in PGPs, as opposed to the traditional fee-for-service system, and to assess the added costs (if any) of allowing the PGP-enrolled Medicaid recipients to retain their Medicaid privileges. Both of the general hypotheses were supported, implying that medical care expenditures for the Medicaid population (where the cost was \$151 per capita per female under 65) can be reduced by enrolling them in PGPs (where the cost was \$124 per capita for females under 65), as long as the PGP-enrolled Medicaid population is not concurrently allowed to retain its Medicaid eligibility (where the cost was \$167 per capita per female under 65). Unfortunately, these data do not allow for any controls in terms of differential health status or the appropriateness of the care received, nor do they consider the out-of-plan use of the PGP sample. Accordingly, Johnson and Azevedo's conclusions are somewhat suspect, although the \$27 to \$43 cost reductions per capita per female under 65 may be of sufficient magnitude to weather any subsequent adjustments for such measurement error.

Other Studies. Christianson and McClure (1979) have examined the competitive effects of HMOs on the delivery of care throughout their geographic area. Specifically focusing on the growth of HMOs in Hawaii (Christianson, 1978) and in the Minneapolis-St. Paul area, where there has been an explosion of HMO growth, Christianson and McClure (1979:812) found that "competition has helped to reduce hospitalization, contain costs and improve access to medical services. At the same time it has focused attention on consumer satisfaction with medical services, increased the range of consumer choice and given consumers better information about providers." Much of this advantageous effect of HMOs on competition (an effect based largely on anecdotal information), however, may be the result of circumstances unique to the Minneapolis-St. Paul situation. Other studies, however, have also demonstrated (to varying degrees) the competitive impact of HMOs on the larger geographic health care system (Lavin, 1978; Goldberg and Greenberg, 1977, 1979; Greenberg, 1977). In addition, Enthoven (1978a, 1978b) has theoretically assessed the potential of HMOs for systematic cost containment, and has reached a favorable conclusion for certain market situations. Accordingly, Christianson and McClure have tempered their earlier statements on the

competitive effects of HMOs by noting that such effective competition is *most likely* to develop in communities meeting three conditions: 1) at least two reputable HMOs are available to the public; 2) employers contribute the same fixed amount for an employee, regardless of the plan the employee chooses, and the selection of cheaper plans brings a rebate to the employee (for the difference); and 3) efficient plans are allowed to increase their membership by offering better coverage at lower premiums.

Greenfield et al. (1978) set out to assess the extent of out-of-plan use of services by Medicare members enrolled in the Health Insurance Plan (HIP). Using data from HIP records and Social Security Administration files, they drew a 10 percent sample (N = 5,202) of HIP Medicare enrollees and a 0.5 percent sample (N = 4,548) of non-HIP Medicare beneficiaries living in the same area. They found that almost one-third of the HIP Medicare enrollees used out-of-plan services, which accounted for the bulk of the \$47 per capita extra that it cost the Medicare system for their care as opposed to the general Medicare population. Thus, enrollment in an HMO (HIP) did not reduce, but actually increased the overall per capita cost of medical care. These results, however, may not be generalized because under Medicare the HIP enrollees were not at risk, even if they used out-of-plan services (because all bills were paid by Medicare anyway). Therefore, patient incentives for using only in-plan services did not exist, nor did HIP exert much pressure to minimize out-of-plan service utilization.

In comparing the hospital cost experience of three competing HMOs, Gavett and Smith (1978:328) focused on elaborating measures of hospital cost experiences. They argued that five factors affect these experiences: "1) the hospital service mix used, 2) the admissions rate, 3) the hospital mix used, 4) the length of stay, and 5) the intensity of resources used within a particular hospital service." Therefore each of these five factors must be taken into consideration when hospitalization costs are examined. Using data from an IPA, a centralized and fully integrated PGP, and a decentralized loosely integrated PGP (and controlling for these five factors), Gavett and Smith (1978:334) concluded that "the difference in risks selected into each of these plans, rather than the structure and controls within these plans and their financial incentives appear to explain most of the differences in hos-

pitalization rates. Those faced with the likelihood of more serious health problems tended to elect the foundation plan, and those less likely to face serious problems elected one of the closed panel plans.”

These data indicate that a modified version of the adverse selection problem is in effect, in which the most vulnerable enrollees choose IPAs in order to continue their existing patient-practitioner relationships, whereas the less vulnerable (who are less likely to have an ongoing patient-practitioner relationship to maintain) choose a PGP for purely economic reasons. This supports the Berki et al. (1977b) data presented earlier. Moreover, Gavett and Smith report that although HMOs reduce hospital costs, the *degree* of cost reduction (after controlling for the five factors) is not as high as the level generally attributed to HMOs. They are quick to point out, however, that the low level of hospital cost savings in their HMOs is probably the result of serious adverse selection, a substantially lower ratio of hospital beds to population in the community, compared with the national average, and the fact that the plan physicians were not effectively at risk for hospitalization costs. These data, then, suggest that traditionally successful HMO performances *may* be artifacts both of measurement error (not controlling for the five factors influencing hospital costs) and of adverse self-selection (comparison group problems), although they do not provide a definitive test of the plan’s incentive structure.

As part of a larger demonstration project, Fuller et al. (1977) evaluated the utilization, cost, and satisfaction experiences of 834 Medicaid recipients experimentally enrolled in a PGP. Data were collected on the experimental group for the 22 months preceding, and the 22 months following their enrollment in the PGP. Data were collected on the Medicaid control group for the first 12 months after the enrollment of the experimental group in the PGP. The compositions of the two groups are strikingly similar. Analysis of the data (Fuller et al., 1977:705) indicates that for those enrolled in the PGP “ambulatory physician encounter rates decreased 15 percent, drug utilization was down 18 percent, hospital admissions decreased 30 percent, and hospital days declined 32 percent after enrollment.”

When overall cost comparisons are made for a comparable benefit configuration (a dental rider was added to the PGP coverage to attract Medicaid enrollees), a 37 percent savings was realized over three years in the PGP compared with the Medicaid population. Unlike other Medicaid studies, that by Fuller et al. reported a voluntary dropout

rate of only 2.5 percent, which they cite as indicating increased satisfaction and good acceptance of the care provided in the PGP. The low dropout rate, however, is more likely a function of the attractiveness of the dental rider in the PGP, which would not be available to the traditional Medicaid population. Unfortunately, Fuller et al. were *not* able to partition the reductions in health services utilization across even the broad categories of appropriate versus nonappropriate care. Accordingly, it is not clear exactly what occurred: Did the PGP skimp? Did the control group overserve? Or some of both?

Pett (1979:42), in a seemingly "well-contained, experimental opportunity," examined the differences in hospitalization rates between government enrollees in a staff model HMO and in an IPA. Because the enrollees in both plans were nearly identical (except for ethnic status), and because the cost and range of benefits in both plans were identical, Pett argued that the only explanation for the marked differences in hospitalization rates was the differences in organizational incentives and physician personnel. He found the admissions per year per 1,000 enrollees to be 56.8 in the staff model and 83.2 in the IPA; hospitalization days per year per 1,000 enrollees were 246.3 in the staff model and 381.7 in the IPA. Unfortunately, Pett was not able to accurately partition the extent of the differences due to each of five incentives that differed in the two HMOs: 1) the IPA physicians were motivated to increase admissions, as their incomes were directly related to their level of activity through their fee-for-service reimbursement system; 2) the staff model physicians were subject to organized peer review, but the IPA physicians were not (throughout most of the study); 3) the IPA physicians expressed considerably more resentment at the notion of control through peer review than did the staff model physicians; 4) the IPA delivered its care through a very loose federation of independent practices, and the staff model delivered all of its care at the well-integrated group practice site; and 5) 80.8 percent of the IPA physicians were foreign medical graduates, whereas 75.7 percent of the staff model physicians were trained in the United States. As a result, although Pett was able to show rather conclusively that the IPA had higher hospital utilization, he could not determine which incentives were responsible. Moreover, because he did not control for the appropriateness of hospitalizations, he could not determine whether the difference was due to skimping in the staff model or to overuse in the IPA. Finally, because he did not report any

data on ambulatory care use, he could not determine whether the staff model was substituting outpatient care for inpatient care.

To determine why individuals and families disenroll from HMOs, Wollstadt et al. (1978) analyzed enrollment and disenrollment data (both voluntary and mandatory) for Medicaid enrollees in the East Baltimore Medical Plan. They found that voluntary disenrollment peaked 3.6 months after the plan had opened, while mandatory disenrollment remained relatively constant throughout the study, suggesting that initial enrollees had higher rates of disenrollment than later enrollees. This may indicate start-up difficulties in establishing patient-practitioner relationships en masse, a hypothesis supported by Forthofer and Glasser (1979). They have shown that the use of physician services by new enrollees in a PGP is significantly lower during their first quarter of enrollment, but their use of diagnostic procedures (laboratory and radiological) is rather high, presumably as a function of the need to establish baseline data. On the other hand, Johnson and Azevedo (1979b) have shown that, over four years of continuous enrollment in a Kaiser Plan, a study of 828 low-income enrollees did not reveal any changes in the annual rate of office visits or drug utilization. Similarly, Mullooly and Freeborn (1979) report that in their six-year (with quarterly data points) retrospective cohort study of Kaiser-Permanente of Oregon, not only were there no start-up costs (either to patients or the HMO) for ambulatory care utilization during a cohort's enrollment, but also the utilization rates for high users, nonusers, and average users all remained relatively constant throughout the study period.

Taken together, these data seem to indicate that although there may not be any start-up costs for the HMO (in terms of a first enrollment quarter utilization surge), there may be some start-up costs for the enrollee (in terms of obtaining a personalized patient-practitioner relationship). When enrollment and utilization are considered at the same time, the potential start-up costs for the enrollees appear to be higher than those for HMOs. Wollstadt et al. (1978:148) found that "of those in the plan for six months, three-quarters of the continuously enrolled had used plan facilities, while less than one in five of the voluntarily disenrolled sought care from the plan. . . . When comparing those enrollees who used the plan facility at least once, differences in facility use between these three groups [enrolled, voluntarily disenrolled, and mandatorily disenrolled] disappeared." Thus, it

would appear that disenrollment is a function of poor or nonexistent patient-practitioner relationships of the traditionally personal style, problems that would be at a peak during the start-up period. This interpretation, however, is somewhat clouded by the fact that these Medicaid recipients are not faced with the same incentive structure as the normally employed HMO population. That is, there is no cost-saving incentive for the Medicaid enrollee to "suffer through" the HMO and its initial poor patient-practitioner relationship. Medicaid enrollees who disenrolled were given back their Medicaid cards, entitling them to comprehensive health care also without cost.

Using data on the use of ambulatory services in five delivery systems, Dutton (1979) focused on the effects of the gross organizational differences in solo practice, fee-for-service group practice, prepaid group practice, public clinics, and hospital outpatient emergency room departments. Multiple regression analyses of her data (Dutton, 1979:221) showed that "sources used primarily by the poor—hospital outpatient departments, emergency rooms, and public clinics—contained important structural and financial barriers, and had the lowest rates of patient-initiated use. The prepaid system in contrast, maximized patients' access to both preventive and symptomatic care and did not seem to inhibit physician-controlled follow-up care."

As a result, the poor, who are not likely to be HMO enrollees because they are generally not members of the employed groups to whom HMO benefits are offered (except through relatively rare Medicare and Medicaid contracts), are deterred from seeking both preventive and somatic care. At the same time, the fee-for-service system encourages physicians to expand their follow-up services, which is quite an expensive process (relatively speaking) for the poor who must use it.

Summary and Conclusions

Two Common Themes

Although the data and conclusions reviewed in this paper were quite varied and often contradicted each other, two common themes can be identified, which future HMO performance studies should take into consideration. *First*, the research designs employed have not been

framed clearly enough, in varying degrees, to provide conclusive evidence to support or reject the hypothesized effects of the incentive and disincentive structures. That is, even in the more sophisticated designs, some complications exist, either because they were not anticipated, or because adjusting the research design to compensate for them was impractical. *Second*, the orientation of these studies has been far more descriptive than analytic (in the deductive, hypothesis-testing sense). Although this tendency is typical for a developing substantive area, a continuation of this trend would be wasteful. Unfortunately, just such a pursuit has recently been suggested for future research on HMOs (see Hester, 1979).

What We Can Say About HMOs

As a result of these two common themes, very little can be conclusively stated about the performance of HMOs. We can, however, make four statements: 1) The total cost for the delivery of health care in an HMO is less than that in a conventionally insured delivery system. 2) The total cost for the delivery of health care in a PGP appears to be less than that in an IPA. 3) The major factor involved in reducing the costs reflected in statements 1 and 2 is the lower level of hospitalization usually found in the HMO (especially in the PGP). 4) *We do not know how or why* statements 1, 2, and 3 are true.

Methodological Deficiencies in the HMO Literature

In addition to these four statements on HMO performance, we may also identify the three most serious deficiencies in the literature of HMO performance. *First*, there is *the design or biased comparisons problem*. That is, in addition to the absence of randomized controlled experimental designs, unadjusted and/or partial comparisons between PGPs, IPAs, and control (conventionally insured) groups are tenuous, at best. This problem may be either blatant (as in the total absence of a control group), or subtle (as in adverse selection). *Second*, there is *the problem of reciprocal causation*. That is, while a study may, for example, focus and report performance data on physician utilization, it may neglect to consider the effects of that physician utilization on hospitalization experiences, and vice versa. In short, the performance of

HMOs on one aspect of the use of services is statistically considered in isolation from the other aspects of service utilization, although they are considered to be related in the conceptual model. *Third*, there is *the problem of reliability and validity of measurement*. For example, when hospitalization is the performance indicator under examination, only the gross rate is measured, rather than the five factors that comprise gross hospitalization experiences.

Emergent Hypotheses

Finally, although the data do not permit us to make any more definite statements about the performance of HMOs, four additional hypotheses have emerged that warrant future study. *First*, those who are already ill apparently pick the IPA (in multiple-choice situations) because they already have a private physician as their regular source of care, as a result of their being ill. That is, if you are already ill, the chances of your having seen a physician on a regular basis (concerning your illness) are higher than if you are not ill. When presented with a multiple-choice situation, it would be more reasonable for you to choose the IPA so that you could continue your patient-practitioner relationship. *Second*, there may well be an anomaly concerning the quality of care received in HMOs. Although PGPs may be more likely to skimp on surgical care than IPAs (or, it would be assumed, than conventional health care systems), the quality of care actually received in the PGP is better. This may reflect the difficulty of simultaneously attaining the HMO's tandem goals of comprehensive health care and cost containment. *Third*, disenrollment from the HMO, among non-Medicaid and non-Medicare enrollees in PGPs, may be a function of the lower quality and quantity of the traditionally personal patient-practitioner relationship to which they have been accustomed. This would be especially critical during initial enrollment periods when patients' demand for personal relationships may considerably exceed the providers' potential supply. *Finally*, enrollment in the HMO may not produce equity in the use of preventive health services; rather, the use of preventive health services may continue to be a function of social class characteristics (especially education and occupational prestige). That is, if, as Luft (1978b) has argued, HMOs do not provide a more preventive orientation than the fee-for-service system, then the effects of education and occupational prestige on preventive service

utilization will not be diluted simply by exposing individuals to HMOs. On the other hand, if HMOs really do have more of a health maintenance orientation (i.e., if Luft is wrong), then the effects of education and occupational prestige on preventive service utilization should be diluted with increased and continued exposure to HMOs.

References

- Aday, L.A., and Andersen, R. 1975. *Access to Medical Care*. Ann Arbor: Health Administration Press.
- , Andersen, R., and Fleming, G., 1980. *Health Care in the United States: Equitable for Whom?* Beverly Hills: Sage Publications.
- Andersen, R. 1968. *A Behavioral Model of Families' Use of Health Services*. Chicago: Center for Health Administration Studies.
- , and Anderson, O. 1979. Trends in the Use of Health Services. In Freeman, H.E., Levine, S., and Reeder, L., eds., *Handbook of Medical Sociology*, 3d edition. Englewood Cliffs: Prentice-Hall.
- , and Newman, J. 1973. Societal and Individual Determinants of Medical Care Utilization in the United States. *Milbank Memorial Fund Quarterly/Health and Society* 51 (Winter):95–124.
- Anderson, O.W., and Sheatsley, P.B. 1959. *Comprehensive Medical Insurance: A Study of Costs, Use and Attitudes Under Two Plans*. New York: Health Information Fund.
- Ashcraft, M., Penchansky, R., Berki, S.E., Fortus, R., and Gray, J. 1978. Expectations and Experience of HMO Enrollees after One Year: An Analysis of Satisfaction, Utilization, and Costs. *Medical Care* 16:14–32.
- Bashshur, R.L., and Metzner, C.A. 1970. Vulnerability to Risk and Awareness of Dual Choice of Health Insurance Plan. *Health Services Research* 5:106–113.
- Bashshur, R., Metzner, C., and Warden, C. 1967. Consumer Satisfaction with Group Practice: The CHA Case. *American Journal of Public Health* 57:1991–1999.
- Berki, S.E. and Ashcraft, M.L.F. 1980. HMO Enrollment: Who Joins What and Why: A Review of the Literature. *Milbank Memorial Fund Quarterly/Health and Society* 58 (Fall):588–632.
- , Ashcraft, M., Penchansky, R., and Fortus, R. 1977a. Health Concern, HMO Enrollment, and Preventive Care Use. *Journal of Community Health* 3:3–31.

- , Ashcraft, M., Penchansky, R., and Fortus, R. 1977b. Enrollment Choice in a Multi-HMO Setting: The Roles of Health Risk, Financial Vulnerability, and Access to Care. *Medical Care* 15:95-114.
- , Penchansky, R., Fortus, R., and Ashcraft, M. 1978. Enrollment Choices in Different Types of HMOs: A Multivariate Analysis. *Medical Care* 16:682-697.
- Broida, J.H., Lerner, M., and Lohrenz, F.N. 1975. Impact of Membership in an Enrolled, Prepaid Population on Utilization of Health Services in Group Practice. *New England Journal of Medicine* 292:780-783.
- Chamberlain, J. 1967. Selected Data on Group Practice Prepayment Plan Services. *Group Health and Welfare News* 8, special supplement: i-viii.
- Christianson, J.B. 1978. *Do HMOs Stimulate Beneficial Competition?* Excelsior, Minn.: InterStudy.
- , and McClure, W. 1979. Competition in the Delivery of Medical Care. *New England Journal of Medicine* 296:812-818.
- Corbin, M., and Krute, A. 1975. Some Aspects of Medicare Experience with Group Practice Prepayment Plans. *Social Security Bulletin* 38:3-11.
- Densen, P.M., Balamuth, E., and Shapiro, S. 1958. *Prepaid Medical Care and Hospital Utilization*. Chicago: American Hospital Association.
- , Jones, E., Balamuth, E., and Shapiro, S. 1960. Prepaid Medical Care and Hospital Utilization in a Dual Choice Situation. *American Journal of Public Health* 50:1710-1726.
- , Shapiro, S., Jones, E.W., and Baldinger, I. 1962. Prepaid Medical Care and Hospital Utilization. *Hospitals* 36:62-68, 138.
- Department of Health, Education, and Welfare. 1974. *Federal Register*, Vol. 38, No. 203. Washington, D.C.: Government Printing Office.
- . 1975. *Health Maintenance Organizations: Guidelines for Subparts B, C, D, E, and G*. Washington, D.C.: Government Printing Office.
- . 1979. *National HMO Census of Prepaid Plans, 1978*. Washington, D.C.: Government Printing Office.
- Diehr, P.K., Richardson, W.C., Drucker, W.L., Shortell, S.M., and LoGerfo, J.P. 1976. *The Seattle Prepaid Health Care Project: Comparison of Health Services Delivery. Chapter 2: Utilization: Ambulatory and Hospital*. Seattle: School of Public Health and Community Medicine, University of Washington.
- , Williams, S.J., Shortell, S.M., Richardson, W.C., and

- Drucker, W.L. 1979a. The Relationship between Utilization of Mental and Somatic Health Services among Low Income Enrollees in Two Provider Plans. *Medical Care* 17:937-952.
- , Richardson, W.C., Shortell, S.M., and LoGerfo, J.P. 1979b. Increased Access to Medical Care: The Impact on Health. *Medical Care* 17:989-999.
- Donabedian, A. 1969. An Evaluation of Prepaid Group Practice. *Inquiry* 6:3-27.
- Dozier, D., Krupp, M., Melinkoff, S., Shwarberg, C., and Watts, M. 1964. *Report of the Medical and Hospital Advisory Council to the Board of Administration of the California State Employees Retirement System*. Sacramento: The State of California.
- Dutton, D.B. 1978. Explaining the Low Use of Health Services by the Poor: Costs, Attitudes, or Delivery Systems. *American Sociological Review* 43:348-368.
- . 1979. Patterns of Ambulatory Health Care in Five Different Delivery Systems. *Medical Care* 17:221-243.
- Ellwood, P. 1971. Health Maintenance Organizations, Concept and Strategy. *Hospitals* 45:53-56.
- Enright, S.B. 1979. I.P.A.: The Initials That May Mean Your Future. *Medical Economics*, February 19, 124-137.
- Enthoven, A.C. 1978a. Competition of Alternative Health Care Delivery Systems. In Greenburg, W., ed., *Competition in the Health Care Sector: Past, Present, and Future*. Washington, D.C.: Federal Trade Commission, Bureau of Economics.
- . 1978b. Consumer-Choice Health Plan. *New England Journal of Medicine* 298:650-658 and 298:709-720.
- Falk, I.S., and Senturia, J. 1960. *Medical Care Program for Steelworkers and Their Families*. Pittsburgh: United Steelworkers of America.
- Forthofer, R.N., and Glasser, J.H. 1979. Utilization of Services of an HMO by New Enrollees. *American Journal of Public Health* 69:1127-1131.
- Freeborn, D.K., Pope, C.R., Davis, M.A., and Mullooly, J.P. 1977. Health Status, Socioeconomic Status, and Utilization of Outpatient Services for Members of a Prepaid Group Practice. *Medical Care* 15:115-128.
- Freidson, E. 1961. *Patients' Views of Medical Practice*. New York: Russell Sage Foundation.
- Fuchs, V. 1974. *Who Shall Live? Health, Economics, and Social Change*. New York: Basic Books.
- Fuller, N.A., Patera, M.W., and Koziol, K. 1977. Medicaid Utilization of Services in a Prepaid Group Practice Health Plan. *Medical Care* 15:705-740.

- Gaus, C.R., Cooper, B.S., and Hirschman, C.G. 1976. Contrasts in HMO and Fee-for-Service Performance. *Social Security Bulletin* 39:3-14.
- Gavett, J.W., and Smith, D.B. 1978. A Comparison of the Hospital Cost Experience of Three Competing HMOs. *Inquiry* 15:327-335.
- Goldberg, L., and Greenburg, W. 1977. *The Health Maintenance Organization and Its Effects on Competition*. Washington, D.C.: Federal Trade Commission, Bureau of Economics.
- . 1979. The Competitive Response of Blue Cross and Blue Shield to the Health Maintenance Organization in Northern California and Hawaii. *Medical Care* 17:1019-1028.
- Goldensohn, S.S., and Fink, R. 1979. Mental Health Services for Medicaid Enrollees in a Prepaid Group Practice Plan. *American Journal of Psychiatry* 136:160-164.
- Goldman, M., and Steinwald, B. 1973. Prepaid Medical Care and Health Services Utilization: A Reexamination of the Evidence. Unpublished.
- Greenburg, W. 1977. HMOs Stimulate Competition, FTC Concludes. *Hospital Progress* 58:10-11.
- Greenfield, C., Densen, P., Jones, E., Kleinman, J., and Miller, J. 1978. Use of Out-of-Plan Services by Medicaid Members of HIP. *Health Services Research* 13:243-260.
- Greenlick, M.R. 1972. The Impact of Prepaid Group Practice on American Medical Care: A Critical Evaluation. *Annals of the American Academy of Political and Social Science* 399:100-113.
- Hastings, J., Mott, F., Barclay, A., and Hewitt, D. 1973. Prepaid Group Practice in Sault Ste. Marie, Ontario. Part 1, *Medical Care* 11:91-103; Part 2, *Medical Care* 11:173-188.
- Hester, J.A. 1979. Research in Resource Allocation in a Prepaid Group Practice. *Milbank Memorial Fund Quarterly/Health and Society* 57 (Summer):388-411.
- Hetherington, R., Hopkins, C., and Roemer, M. 1975. *Health Insurance Plans: Promise and Performance*. New York: Wiley.
- Johnson, R.E., and Azevedo, D.J. 1979a. Examining the Annual Drug Utilization of a Cohort of Low Income Health Care Plan Members. *Medical Care* 17:578-591.
- . 1979b. Comparing the Medical Utilization and Expenditures of Low Income Health Plan Enrollees with Medicaid Recipients and Low Income Enrollees Having Medicaid Eligibility. *Medical Care* 17:953-966.
- Klarman, H.E. 1963. Effect of Prepaid Group Practice on Hospital Use. *Public Health Reports* 78:955-965.

- Kronenfeld, J.J. 1978. Provider Variables and the Utilization of Ambulatory Care Services. *Journal of Health and Social Behavior* 19:68-76.
- Lairson, D.R., and Swint, J.M. 1978. A Multivariate Analysis of the Likelihood and Volume of Preventive Visit Demand in a Prepaid Group Practice. *Medical Care* 16:730-739.
- . 1979. Estimates of Preventive versus Nonpreventive Medical Care in an HMO. *Health Services Research* 14:33-43.
- Lavin, J.B. 1978. The Town That's Squeezing Out Private Practice. *Medical Economics*: July 24, 27-44.
- Levin, B.L., and Glasser, J.H. 1979. A Survey of Mental Health Service Coverage within Health Maintenance Organizations. *American Journal of Public Health* 69:1120-1125.
- LoGerfo, J.P., 1975. Hypertension: Management in a Prepaid Health Care Project. *Journal of the American Medical Association* 233:245-248.
- , Lawson, E., Okamoto, G., Diehr, P.K., and Richardson, W. 1977. Quality of Care for Common Infections: Comparisons of Prepaid Group Practice and Independent Practice. Paper presented at the meeting of the American Federation for Clinical Research.
- , Lawson, E., and Richardson, W.C. 1978. Assessing the Quality of Care for Urinary Tract Infection in Office Practice: A Comparative Organizational Study. *Medical Care* 16:488-495.
- , Efrid, R.A., Diehr, P.K., and Richardson, W.C. 1979. Rates of Surgical Care in Prepaid Group Practices and the Independent Setting: What Are the Reasons for the Differences? *Medical Care* 17:1-11.
- Luft, H.S. 1978a. How Do Health Maintenance Organizations Achieve Their Savings? Rhetoric and Evidence. *New England Journal of Medicine* 298:1336-1343.
- . 1978b. Why Do HMOs Seem to Provide More Health Maintenance Services? *Milbank Memorial Fund Quarterly/Health and Society* 56 (Spring):140-168.
- . 1979. HMOs, Competition, Cost Containment, and NHI. Paper presented at the meetings of the American Enterprise Institute, Washington, D.C. October.
- . 1980a. Trends in Medical Care Costs: Do HMOs Lower the Rate of Growth? *Medical Care* 18:1-16.
- . 1980b. *Health Maintenance Organizations: Dimensions of Performance*. New York: Wiley.
- Mechanic, D. 1979. Correlates of Psychological Distress among Young Adults: A Theoretical Hypothesis and Results from a

- 16-Year Follow-up Study. *Archives of General Psychiatry* 36:1233-1239.
- . 1980. The Experience and Reporting of Common Physical Complaints. *Journal of Health and Social Behavior* 21:146-156.
- Mullooly, J.P., and Freeborn, D.K. 1979. The Effect of Length of Membership upon the Utilization of Ambulatory Care Services: A Comparison of Disadvantaged and General Membership Populations in a Prepaid Group Practice. *Medical Care* 17:922-936.
- National Advisory Commission. 1967. *Report of the National Advisory Commission on Health Manpower*. Volume 11:197-228. Washington, D.C.: Government Printing Office.
- Nixon, Richard. 1971. *Health Message of the President of the United States*. Washington, D.C.: Government Printing Office.
- Patrick, D.L., Coleman, J.V., Eagle, J., and Nelson, E. 1978. Chronic Emotional Problem Patients and Their Families in an HMO. *Inquiry* 15:166-180.
- Perrott, G. 1971. *The Federal Employees Health Benefits Program: Utilization Study*. Washington, D.C.: Health Services and Mental Health Administration.
- , and Chase, J.C. 1968. The Federal Employees Health Benefits Program: Sixth Term Coverage and Utilization. *Group Health and Welfare News*, special 8-page supplement, October.
- Pett, L.J. 1979. Utilization of Hospital Services in the Individual Practice Association on Guam: A Comparative Analysis. *Medical Care* 17:420-429.
- PL 93-222. 1973. *The Health Maintenance Organization Act of 1973*. Washington, D.C.: Government Printing Office.
- PL 93-641. 1974. *The National Health Planning and Resources Development Act of 1974*. Washington, D.C.: Government Printing Office.
- PL 94-460. 1976. *Health Maintenance Organization Amendments of 1976*. Washington, D.C.: Government Printing Office.
- PL 95-559. 1978. *Health Maintenance Organization Amendments of 1978*. Washington, D.C.: Government Printing Office.
- Pope, C.R. 1978. Consumer Satisfaction in a Health Maintenance Organization. *Journal of Health and Social Behavior* 19:291-303.
- Riedel, D., Walden, D., Singesen, A., et al. 1975. *Federal Employees Health Benefits Program: Utilization Study*. Washington, D.C.: National Center for Health Services Research.
- Robertson, R.L. 1972. Comparative Medical Care Use Under Prepaid Group Practice and Free Choice Plans: A Case Study. *Inquiry* 9:70-76.

- Roemer, M.J., and Shonick, W. 1973. HMO Performance: The Recent Evidence. *Milbank Memorial Fund Quarterly/Health and Society* 51 (Spring):271-317.
- Scitovsky, A.A., Benham, L., and McCall, N. 1979. Use of Physician Services under Two Prepaid Plans. *Medical Care* 17:441-460.
- , McCall, N., and Benham, L. 1978. Factors Affecting the Choice between Two Prepaid Plans. *Medical Care* 16:660-681.
- Shapiro, S., Williams, J., Yerby, A., Densen, P., and Rosner, H. 1967. Patterns of Medical Use by the Indigent Aged under Two Systems of Medical Care. *American Journal of Public Health* 57:784-790.
- Shortell, S., Richardson, W., LoGerfo, J., Diehr, P.K., Weaver, B., and Green, K. 1977. The Relationships among Dimensions of Health Services in Two Provider Systems: A Causal Model Approach. *Journal of Health and Social Behavior* 18:139-159.
- Weinerman, E.R. 1964. Patients' Perceptions of Group Medical Care: A Review and Analysis of Studies of Choice and Utilization of Prepaid Group Practice Plans. *American Journal of Public Health* 54:880-899.
- Wersinger, R., Roghmann, K.J., and Gavett, J.W. 1976. Inpatient Hospital Utilization in Three Prepaid Comprehensive Health Care Plans Compared with a Regular Blue Cross Plan. *Medical Care* 14:721-732.
- Williams, J.J., Trussell, R.E., and Elinson, J. 1964. A Survey of Family Medical Care under Three Types of Health Insurance. *Journal of Chronic Diseases* 17:879-884.
- Williams, S.J., Shortell, S.M., LoGerfo, J.P., and Richardson, W.C. 1978. A Causal Model of Health Services for Diabetic Patients. *Medical Care* 16:313-326.
- , Diehr, P.K., Drucker, W.L., and Richardson, W.C. 1979. Mental Health Services: Utilization by Low Income Enrollees in a Prepaid Group Practice Plan. *Medical Care* 17:139-151.
- Williamson, J.W., Cunningham, F.C., and Ward, D.J. 1979. Quality of Health Care in HMOs as Compared to Other Health Settings: A Literature Review and Policy Analysis. Paper prepared for the Office of Health Maintenance Organizations, Department of Health, Education, and Welfare, Rockville, Md.
- Wolinsky, F.D. 1976. Health Service Utilization and Attitudes toward Health Maintenance Organizations: A Theoretical and Methodological Discussion. *Journal of Health and Social Behavior* 17:221-236.
- . 1978. Assessing the Effects of Predisposing, Enabling, and Illness-Morbidity Characteristics on Health Service Utilization. *Journal of Health and Social Behavior* 19:384-396.

———. 1980. *The Sociology of Health: Principles, Professions, and Issues*. Boston: Little, Brown.

Wollstadt, L.J., Shapiro, S., and Bice, T.W. 1978. Disenrollment from a Prepaid Group Practice: An Actuarial and Demographic Description. *Inquiry* 15:142-149.

The author expresses his appreciation to Leonard A. Ferber for his assistance in preparing the bibliography, to Louis J. Goodman for his comments and suggestions, and to Lynn E. Jensen for supporting this project. Suggestions of the reviewers have significantly improved this paper. The views expressed here are the author's and do not necessarily represent the official position or policy of the American Medical Association.

Address correspondence to: Fredric D. Wolinsky, Ph.D., Department of Statistical Analysis, Center for Health Services Research and Development, American Medical Association, 535 N. Dearborn, Chicago, IL 60610.