Health Maintenance Organizations

DAVID STRANG

By Harold Luft's (1978, p. 1336) definition, a health maintenance organization (HMO) "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 sources of revenue." HMOs are best contrasted with the fee-for-service, third-party payment arrangements that dominate American medicine, where an insurer reimburses an independent provider (or indemnifies the insured customer) for services rendered. A short organizational understanding of an HMO is that it combines the health insurance and health delivery functions generally kept separate in American medicine.

Health maintenance organizations thus make up a population of firms set off organizationally, though not technologically, from the larger health industry or medical care system. HMOs are of particular interest from an organizational perspective. They involve an effort to integrate, and thus potentially to manage, the provision of health care. This runs counter to the conventional structure of health-care delivery, where increasingly sophisticated work is characteristically handled not through complexity in organizational design and strategy, but through the complexity of the physician as an autonomous professional. In effect, health maintenance organizations embody a fundamental organizational innovation in health care.

Health maintenance organizations are not only of interest to organizational researchers, however. Health consumers, doctors, purchasers, and policymakers have all seen the HMO as having great promise as a vehicle for consumer control, for expanded collegiality among doctors, and for reducing health costs. At the same time, HMOs have been highly controversial, due to the challenge they pose to the professional autonomy of the physician. With government support in the 1970s and 1980s, the prepaid group movement matured into the HMO industry. HMOs form an increasingly substantial part of the American
health-care system, and one with the potential to reorganize standard ways in which medicine is financed and delivered.

TWO HMO PORTRAITS

The leading HMO throughout the post–World War II period has been Kaiser-Permanente. Formed during World War II, Kaiser enrollments topped one million in 1963. In the early 1970s, the Kaiser plans enrolled about half of all HMO members (over two million) in the country. While that figure is now about 15 percent, Kaiser remains by far the largest single HMO (or HMO chain). Further, Kaiser-Permanente has long been regarded not only as one of the biggest HMOs, but one of the best organized. Kaiser-Permanente plans compare favorably to both fee-for-service medicine and other HMOs on a variety of dimensions: controlling costs, maintaining quality of care, providing access and generating patient and provider satisfaction (for a review, see Somers 1971).

The Kaiser plans began not with Henry Kaiserbut with Dr. Sidney Garfield, a physician who in 1933 set up a hospital to care for construction workers in the California desert. Unable to maintain a fee-for-service practice in such a remote location, Garfield started a prepaid health plan. Henry J. Kaiser, a prominent industrialist who had managed some of the California construction sites, then recruited Garfield to set up a similar plan in a remote construction site in Washington. In both places, prepayment proved the best way to organize health care to concentrated groups in areas lacking an existing medical infrastructure.

During World War II, Kaiser and Garfield entered into a closer relationship to provide health care for Kaiser shipbuilding employees in Oakland, California. It was with this practice that Kaiser-sponsored prepaid practice first became a permanent organization. After World War II, Kaiser took formal control of the plan’s hospitals and other services, while Garfield’s practice became the Permanente Medical Group. The plan began to provide health care not only to Kaiser employees, but to other large corporations as well.

Kaiser-Permanente faced strong opposition from the local medical society, which accused Garfield of advertising and soliciting patients, preventing patients from having a free choice of physicians, rendering inadequate services, and channeling profits into his health plan. But unlike most HMOs, strong financial backing received from the Kaiser industries enabled the Kaiser plans to set up a wholly autonomous medical system. Kaiser-Permanente built its own hospitals and staffed them with a dedicated medical group (i.e., Kaiser hospitals were staffed only by Permanente physicians, and Permanente physicians worked only in Kaiser hospitals). Kaiser set up an in-house research division, and even considered establishing its own medical school. The autonomy of the Kaiser system made it relatively invulnerable to the occupational pressures that overcame many prepaid health plans between the 1940s and the 1960s.

As in many prepaid plans, expansion in size led to tension between Kaiser and the physicians staffing the organization. Physicians resented Henry Kai-

ser’s control over plan expansion, while Kaiser opposed physician initiatives and the threat that Permanente physicians might contract with competing plans. In 1955, Kaiser-Permanente was reorganized to more clearly distinguish areas of authority between the hospital and marketing arms of the plan (run by Kaiser executives) and the health delivery arm (run by Permanente medical groups) on the other. This dual management structure proved highly effective in promoting a strong organizational culture of cooperation between administrators and physicians. It also facilitated the later expansion of Kaiser-Permanente, which was able to organize separate plans on the same model in Hawaii, Colorado, Cleveland, and Washington, D.C.

While Kaiser-Permanente is the outstanding example, a similar story of internal integration, organizational innovation, and steady growth could be told for a number of HMO prototypes: Group Health Association in Washington, D.C., Health Insurance Plan of New York, Group Health Cooperative of Puget Sound. Each of these plans combined some form of entrepreneurial or consumer management with a strong physician staff, faced tensions over plan threats to the professional autonomy of physicians, and provided economical health care to large employee groups in major urban centers.

A more interesting contrast is to United Healthcare (see Moore 1979; Moore, Martin, and Richardson 1983; Martin, Ehret, and Geving 1983). In the early 1970s, SAFECO Insurance Company began to consider developing a prepaid health care plan as a part of its insurance offerings. Its interest in doing so was spurred by the shift toward federal support for HMOs (to be discussed). SAFECO sought to make use of its existing network of marketing representatives to promote its medical insurance plan, while contracting with individual physicians and group practices to provide health care.

Between 1974 and 1979, SAFECO sponsored several plans that were eventually combined to form United Healthcare (UHC). United Healthcare contracted mainly with individual primary-care physicians. It was widely noted for making early use of a “gatekeeper” structure, where primary-care physicians controlled referrals to specialists (that is, patients could not self-refer). The plan reimbursed gatekeepers on the basis of their standard fee-for-service charges. Ten percent of physician charges was withheld, to be returned to the physician if his or her total costs per patient were below a target figure.

This structure facilitated rapid growth, with both physicians and members rather easily signed up. By 1978, UHC had expanded into Woodland, California; Seattle, Spokane, and Bellingham, Washington; and Salt Lake City, Utah. The plan grew at a rate of 75 percent per year throughout the 1970s, peaking at about 38,000 subscribers in 1980. In 1979, optimistic discussions of United Healthcare and the advantages of a gatekeeper-based system were reported in the New England Journal of Medicine and the Wall Street Journal.

But United Healthcare costs per enrollee mounted along with the expansion of the plan. United Healthcare had failed to select cost-conscious providers, instead attempting to contract with most physicians in covered areas so prospective enrollees would not have to change doctors. Reimbursement of physicians on a fee-for-service basis with a small withheld amount proved insufficient to induce economizing behavior among physicians.2 UHC had no controls
over specialists, no system for monitoring the physician's utilization of services or certifying hospital stay, and no program for educating physicians about cost-conscious styles of practice.

In 1981, United Healthcare sought to restructure its relation to physicians. Panels of specialists were contracted to bring referral costs under control, high-cost doctors were dropped from the plan, and the size of withheld provider fees was doubled. While costs began to decrease with these organizational reforms, SAFECO sought a buyer for United Healthcare. When several possible purchases fell through, SAFECO terminated United Healthcare in 1982.

The contrast between Kaiser-Permanente and United Healthcare suggests the importance of several features of health maintenance organizations. First, HMOs involve complex relations between groups of physicians and external sponsors (even in those cases where physicians themselves initially sponsor the HMO, a distinction develops over time between the physicians and the plan). Second, the internal structure of the plan is crucial for HMO survival and profitability; particular importance are modes of reimbursement, relations between physicians, and managerial monitoring and education. Third, the environmental conditions under which HMOs develop have changed dramatically over time, with substantial consequences for organizational strategies and success.

HISTORICAL BEGINNINGS

While the term health maintenance organization only gained currency in the 1970s, prepaid practice is a fairly common mode of organizing medicine. Abel-Smith (1988) notes that it was common practice in many European countries before World War II for occupational or social groups to hire physicians to provide medical care for a prepaid fee. Similarly, in the first decades of the twentieth century, company doctors reimbursed on a salaried or prepaid basis were widespread in the railroad, mining, and lumber industries (Starr 1982, p. 203). But these forms of "contract medicine" have largely vanished, due to fierce opposition from increasingly united medical professions (and in Europe, from the implementation of national health insurance).

The roots of the contemporary HMO industry can be located in health cooperatives formed in the 1920s and 1930s. Of particular importance was Dr. Michael Shadid's medical cooperative at Elk City, Oklahoma. Dr. Shadid sought to provide comprehensive and up-to-date medical care in a small rural community by setting up a cooperatively organized prepaid health plan. While colleagues rejected Shadid's idea and sought to deprive him of his medical license, the Oklahoma Farmers Union supported the plan. Fierce opposition from the local medical community eventually led to the dissolution of the Elk City Cooperative. But Shadid lectured widely about the promise of health cooperatives and succeeded in inspiring a number of companion efforts.

Many of the early HMOs (or prepaid group practices, as they were then called) flowed from the aim to increase consumer control and medical cooper-

HEALTH MAINTENANCE ORGANIZATIONS

ation. Labor unions provided an important early impetus through their efforts to make medical care available to workers in isolated company towns. Cooperative designs were realized successfully by Group Health Association, a consumer cooperative begun in 1938 by a group of federal employees, and Group Health Cooperative of Puget Sound, a physician collective founded in 1947. HMO "prototypes" were also founded by corporate rationalizers aiming for more efficient and accessible healthcare. The most important corporate rationalizer was Henry J. Kaiser, as detailed previously. Similarly, Fiorello La Guardia set up Health Insurance Plan of New York in 1947 to serve much of the city's public workforce. While early prepaid plans responded to various perceptions of the inadequacies of conventional arrangements, they shared much pressure from the medical community. As Paul Starr (1982) details, doctors have vigorously opposed the emergence of organized clients (including the development of health insurance), the imposition of organizational limits on professional autonomy, and the outbreak of price competition among providers. In 1934, for example, the American Medical Association (AMA) described all medical institutions as "but expansions of the equipment of the physician" and insisted that "no third party must be permitted to come between the patient and his physician." The AMA also required that the immediate cost of medical care be borne by the patient, and that health plans should include all physicians who wished to participate in them (Starr 1982, pp. 299-300).

Prepaid plans directly violated these prescriptions, challenging the physician's workplace autonomy and financial control. They could underprice and outcompete solo practitioners, were often sponsored and managed by laypeople, and inserted an organizational structure between physician and patient. In response, the medical profession brought considerable political and organizational pressure to bear on nascent HMOs. Physicians joining HMOs were barred from memberships in county medical societies and admitting privileges in hospitals, threatened with the loss of their medical licenses, and confronted with general social and professional ostracism. Under pressure from organized medicine, many states passed Blue Cross/Blue Shield laws requiring prepaid plans to be sponsored by physicians or gain county medical society approval, to allow all local physicians to participate, and to guarantee physician control of HMO-governing bodies (Starr 1982, pp. 303-6).

Resistance to HMOs also took the form of organization building. In 1956, solo practitioners in the Sacramento area formed the San Joaquin Foundation for Medical Care to resist the threat posed by an expanding Kaiser-Permanente. The San Joaquin Foundation also delivered medical care to prepaid enrollees, but it was run by and for physicians who maintained their existing private practices. While formed as an oppositional device, these medical-care foundations (or independent practice associations, as they later came to be known) had enough in common with the prepaid group plans that subsequent federal and state HMO policy treated the two as alternative forms of HMOs.

The HMO population grew slowly from the 1950s through the 1960. Kaiser-Permanente and some other plans emerged as major providers, primarily on
the West Coast and in some large eastern and midwestern cities. But in general
HMOs failed to diffuse widely. While some HMOs grew throughout the post-
war period, there were on the order of twenty HMOs in the country in 1960,
and less than forty in 1969. The 1960s saw the development of few plans of the
nature of the early health cooperatives and industry-organized plans.

On the political side, conditions improved somewhat for HMOs over the
decades. The courts sometimes blocked the efforts of the medical profession to
restrict prepaid plan access to health facilities. Most notably, Group Health
Association (GHA) won a suit against the Washington, D.C., medical society,
which had attempted to bar GHA physicians from admitting patients to hospi-
tals. And some accommodation between prepaid plans and organized medicine
was reached in 1960, when the AMA recognized free choice of health plan as
a substitute for free choice of physician. But HMOs remained a marginal and
degenerated population within the American health-care system.

Further, the conditions of the early growth of health maintenance organiza-
tions—lack of accessible health care for labor unions and businesses on the one
hand, and the vigour of the cooperative movement on the other—seemed to be
disappearing. When HMOs succeeded in the 1970s and 1980s, it was on the
basis of a new sort of appeal: their promise as vehicles of cost containment in
an era of health inflation.

INSTITUTIONAL LEGITIMATION

The term health maintenance organization itself is tied to this shift. It was
coined in the late 1960s by Dr. Donald Ellwood, executive director of the Ameri-
can Rehabilitation Federation, a health policy analysis group. Ellwood (and
others, notably Alain Enthoven, Jon Christianson, and Clark Havighurst) ad-
vocated prepaid group practice as a cure for the rampant cost inflation that had
accompanied the infusion of federal funds for Medicare and Medicaid. Mount-
ing health costs were traced to misaligned incentives within the health-care
system. Fee-for-service, third-party payment arrangements gave the physician
a free hand to engage in expensive forms of care. On the other hand, the insurer
who footed the bill possessed an incentive to economize on care, but no oppor-
tunity to translate this incentive into policy.

Ellwood argued that an organization that not only insured individuals but
provided their health care would realign incentives and capacities in an appro-
priate way. Such an organization would have the ability to manage the delivery
of health care in an economizing fashion. Ellwood christened prepaid plans as
"health maintenance organizations" to summarize his boldest (and least
broadly supported) claim, that prepaid groups would achieve these goals by
replacing medicine's usual focus on crisis management with attention to preven-
tive care.

Ellwood succeeded in stimulating health bureaucrats at HEW (Department
of Health, Education, and Welfare) to putting the newly named health main-
tenance organizations on the federal agenda. HMOs proved politically attrac-
tive, allowing the Nixon administration to offer a market-based alternative to
National health insurance. And the economic arguments of HMO advocates,
who emphasized the immediacy of financial incentives and minimized the labor
of organizational creation and maintenance, proved highly compelling to health
services experts and public policymakers alike. In 1971, Nixon made HMOs the
centerpiece of his program for health-care reform. Two years later, legis-
lation was passed by Congress and signed into law.

The HMO Act of 1973 overrode restrictive state legislation, instituted a pro-
gram of grants and loans, and required employers to offer HMOs as an option
within their health-care programs. All these provisions applied to "federally
qualified HMOs," however, and leading plans like Kaiser viewed the qualifi-
cation requirements as burdensome. Rather than seeking qualification, major
HMOs campaigned to eliminate the broad benefit packages, unrestricted en-
rollment, and community rating mandated by federal law. Most of these re-
quirements were lifted or weakened in amendments to the act passed in 1976
and 1978 (for excellent legislative histories, see Falkson 1980; Brown 1983).

The states were quick to follow the lead of the federal government. Within
two years of the federal HMO law, twenty-six states had passed enabling acts
to certify and regulate HMOs. These laws superseded earlier statutes impeding
organizational formation, and put HMO regulation on a more positive and less
uncertain basis. Like federal legislation, state laws were not a response to an
increasingly active HMO lobby; in fact, seventeen states without HMOs passed
HMO-enabling legislation. Instead, HMO laws were passed to counter the ris-
ing costs of medicine, with legislation occurring most rapidly and most favora-
ble to HMOs where costs were highest (Strang and Bradburn 1993).

It should be noted that the policy appeal of HMOs remains strong. HMOs
came to be viewed as an instrument for the restructuring of medicine, most promi-
nently in the Clinton administration's aim of grouping providers into large
HMOs competing for the health dollars of large-scale consumer organizations.
This effect draws on the analysis developed by Alain Enthoven and Paul Ell-
wood in the 1970s, where HMOs are seen as capable of integrating and thus
properly aligning financial incentives within the context of a market for medical
care.

POPULATION BOOM AND STABILIZATION

Shifts in the legal and policy environment helped trigger a boom in health main-
tenance organizations. Figure 8-1 plots HMO organizational density (numbers
of operating HMOs) over the 1970s and 1980s, and total HMO enrollment. Sta-
tistics are not readily obtained for industry size prior to the 1970s, since the
various forms of HMOs were not treated as subcategories of a meaningful
larger category until that time. For more extensive reviews of HMO demo-

HMO growth was steady but unimpressive in the decade following the fed-
eral HMO Act. While falling far short of ambitious federal projections that
1,700 HMOs would be in operation by 1975, the number of HMOs increased
fivefold from 1971 to 1982 (from 46 to 262 HMOs). Enrollment also increased
substantially, though not in proportion to growth in organizations (rising from 3.6 million to 10.7 million enrollees).

Industry growth accelerated rapidly in the mid-1980s, sparked by the increasing availability of state Medicare contracts and the development of a highly favorable capital market for HMO start-ups. The number of operating HMOs tripled, reaching some 643 HMOs in 1988. Nearly every major metropolitan market was entered by at least one HMO. HMO enrollment also tripled, increasing from 10 million to over 30 million enrollees.

Rapid industry growth in the 1980s was accompanied by a number of changes in industry composition. In the early 1970s, HMOs were highly concentrated in a few metropolitan areas. In 1975, sixty-eight HMOs with a combined enrollment of 3 million were located in California (making up 40 percent of all plans and 55 percent of all HMO subscribers). Thirty-five of these plans operated in Los Angeles alone. Other centers of HMO development were Minnesota and Wisconsin, while major cities like New York, Chicago, and Boston had few HMOs. And in contrast to the Pacific region, in 1973 the South could boast only five HMOs with a total enrollment of 17,000 subscribers.

By 1988, all states except Alaska were home to at least one HMO. California continued to have more HMOs (51) and HMO enrollment (7 million) than any other state. But now Florida, Illinois, New York, Ohio, Texas, and Wisconsin all had 30 or more operating HMOs. Regional differences were all but invisible.

Figure 8-1  Density and enrollment of health maintenance organizations. (Source: InterStudy, various years.)

As HMOs were formed wherever population density provided substantial markets. Southern states, perhaps the most inhospitable to prepaid practice in the pre-Ellwood era, had one quarter of all operating plans (167 HMOs).

The 1970s and 1980s also saw important shifts in distributions of organizational size and scope. Before the late 1970s, only Kaiser-Permanente had control over a variety of distinct plans located in several states. But attractive financial markets in the 1980s led to a large-scale flowering of national HMO chains. Through acquisitions and new starts, national HMOs came to control 300 HMOs by 1987, or about half of the entire HMO population. Organizationally, HMOs remained highly local operations, since health care is a highly local activity. But financially, and to some extent strategically, HMO development became a national rather than a local phenomenon.

The nationalization of fiscal control occurred without much change in the average size of operating units. Enrollment per HMO dropped in the 1970s, and then held steady in the 1980s at about 40,000 members per HMO. (Organizational size is better captured by the total enrollment of HMOs than by the number of HMO employees, since different HMOs contract for different percentages of the physician’s time.) But most HMOs at any one time have fewer than 15,000 enrollees, while the less than 10 percent of HMOs with more than 100,000 members have about two thirds of total HMO enrollment.

The HMO population suffered a shakeout in the late 1980s, with more than 100 organizational failures in the last three years of the decade and a reduction in population size of about 15 percent. (Prelar years had witnessed failure rates on the order of 10–20 percent of operating plans per year, but these shortages occurred in the context of an overall expansion in the organization’s population.) The favorable capital markets that had fueled rapid growth petered out as HMOs not only filled their niche but overflowed it.

In large part this shakeout resulted directly from the rapid expansion of the previous several years. Many of the plans that started during the boom were the product of easy opportunities for external investors (Moran and Savela 1986), who could profit even if HMOs never got off the ground. The national chain-sponsored IPAs founded during the period of the boom were especially likely to fail (Christianson et al. 1991). This pattern is reminiscent of Hannan and Carroll’s (1992) arguments about “density delay,” where the presence of many similar firms is argued to squeeze newly formed organizations. But for HMOs, it seems even more apparent that periods when many weakly established organizations are carelessly thrown together constitute booms in organizational founding. It is unsurprising that such organizations are failure prone.

**HMO Dynamics and Local Market Conditions**

While most ecological analyses have studied population dynamics in a single site (e.g., Hannan and Freeman 1987) or have analyzed a number of sites in-
DENSITY-DEPENDENT EVOLUTION

dependently (Carroll and Hanan 1989), the fundamentally local operation of most HMOs until the mid-1980s encourages simultaneous multimarket analyses. These analyses permit close investigation into the local conditions that facilitate HMO growth and maintenance that is easily achieved in single time series studies. A number of studies by health services researchers have examined conditions associated with HMO presence and size (Goldberg and Greenberg 1981; Morrissey and Ashby 1982; Welch 1984; McLaughlin 1987). More recently, a number of studies have sought to test ecological arguments about patterns of organizational founding (Strang and Uden-Holman 1990; Wholey, Sanchez and Christianson 1993) and failure (Wholey, Christianson, and Sanchez 1992).

This research develops a consistent portrait of the conditions that lead to HMO founding and maintenance. HMOs are found in the larger metropolitan areas, and in areas where their typical enrollees form large segments of the population (areas with large businesses and much in-migration, families with children, individuals with high income or education). They are more likely to be formed where physicians are easily recruited (where there are many physicians per capita, and where many young physicians are struggling to set up practices). HMO growth is also encouraged by high medical costs, since HMOs are seen as vehicles for cost containment. And finally, HMOs are formed more easily where state legislation is favorable (though Wholey, Sanchez, and Christianson 1993) note that by permitting more weak plans to get off the ground, favorable state regulation may also raise failure rates.

Ecologically oriented studies have focused on the influence of organizational density on founding, and organizational size on mortality. Wholey, Sanchez, and Christianson's (1993) study of founding in the 1980s finds the nonmonotonic patterns of density that Hannan and Freeman (1987) argue should result from opposing effects of population legitimacy and competition. (That is, at low density additional HMOs primarily add to perceptions of the appropriateness of HMOs and the inception of these perceptions in law; at high densities additional HMOs primarily tighten competition for patients and providers.) Strang and Uden-Holman (1999) show that in the 1970s the HMOs that were strongly opposed by the medical community (the prepaid group practice described later in this chapter) exhibit only positive effects of existing HMO presence in a community. Presumably institutional and attitudinal conditions were sufficiently antagonistic that competitive pressures between HMOs were outweighed by the extent to which later HMOs benefitted from the regulatory and occupational struggles waged by the first HMOs to enter the community.

Wholey, Christianson, and Sanchez (1992) show that HMO mortality rates vary with organizational size, though in different ways for different subpopulations. They interpret low rates of failure for plans organized around a multi-specialty group practice or hospital staff as the product of high commitment by cohesive groups of doctors. By contrast, they argue that plans whose physicians do not form integrated groups are likely to experience high rates of failure when small.

HEALTH MAINTENANCE ORGANIZATIONS

THE ORGANIZATIONAL STRUCTURE OF HMOs

While organizational structures and strategies vary widely within the HMO population (typical variations are discussed in the next section), they share a number of basic organizational features. Almost all HMOs employ a dual structure, with separate administrative and provider arms. An executive director heads the administrative side of the organization, which generally includes marketing, financial, personnel, and legal departments or functions (depending on organizational size). A medical director heads the provider side of the organization, typically supported by a variety of physician committees on quality assurance, utilization review, staffing, and the like.

All HMOs face difficulties in organizing and managing the work of the "sovereign" physician (Freidson 1970; Starr 1982). HMO managerial styles emphasize physician education and mutual cooperation rather than discipline, and relations between the plan and the physician involve mutual negotiation rather than control. Nevertheless, HMOs are in the business of containing costs and promoting innovation in physician practice patterns.

Formal management efforts include financial incentives, requirements that the plan authorize some medical decisions (typically hospital admissions and out-of-plan referrals), the collection and dissemination of physician utilization records, and the development of standard practice protocols. Informal efforts include intraphysician socialization to plan norms and advocacy by the medical director and other leaders of physicians of particular practice patterns. While these efforts can add up to very little (Freidson 1975), most research finds that elaborated and integrated managerial systems can successfully mold physician practice patterns (Eisenberg 1986; Fox and Heinen 1987; Hillman, Pauly, and Kerstein 1989; Strang and Currivan 1992).

The bulk of the health services literature on HMOs is devoted to exploring differences between HMOs and fee-for-service arrangements, particularly with reference to health costs. The most comprehensive analysis of HMO cost-containment performance remains Harold Luft's (1981) integration of a large number of prior studies. Luft found that HMO savings result primarily from the substitution of ambulatory care for expensive hospital care. More recent debate has focused on whether HMO savings are due to changes in provider patterns or to selective enrollment or disenrollment (where HMOs mainly serve the healthy). Important support for the former was found in a massive RAND experiment that randomly assigned individuals to a major HMO or fee-for-service medicine (see Manning et al. 1984).

TYPES OF HEALTH MAINTENANCE ORGANIZATIONS

The conventional HMO typology describes four HMO forms: the staff, group, network, and independent practice association (IPA) models. For many purposes these are usually collapsed into two, contrasting the IPA with the other three plan types, often described jointly as prepaid group practices, or PGPs.
It is important to note that these classifications are generated by industry participants, rather than by academic analysts. This detracts from their analytic clarity, but helps ensure their relevance for immediate organizational concerns.

Plan types are differentiated on the basis of the physician's legal and organizational setting. Staff HMOs are plans where most physicians are directly employed by the HMO. Group HMOs are plans where most physicians are members of a single multispecialty group practice contracting with the HMO, and network HMOs are plans where most physicians are members of two or more such group practices. IPAs are plans where most physicians are solo practitioners contracting with the HMO, either directly or through a physician association.

Distinctions between plan types are important for their historical and structural relation to plan-enrollee relations, and even more to relationships internal to the organization. Along almost all relevant dimensions, IPAs are found at one extreme and staff models at the other, with group and network HMOs closer to staff than IPA plans. To simplify, I mainly compare the three sorts of prepaid group practices to IPAs, with the understanding that staff models best exemplify what is distinctive about PGPs.

First, physicians in the different kinds of PGPs tend to see only or mostly HMO enrollees. By contrast, physicians in IPAs typically maintain a large non-HMO practice, with on the order of 10 percent of their patients enrolled in the HMO. This contrast is grounded in the founding rationales of the different kinds of plans. As noted above, IPAs were generally formed by solo practitioners to oppose the inroads of large PGPs. The aim behind the IPA was thus to preserve the private practices of its physicians, not eliminate them.

Second, PGPs are unlikely to reimburse physicians on a fee-for-service basis, whereas IPAs and PGP-IPAs are generally salaried; in group and network models, they are capitated. (Capitation refers to the payment of fee per enrollee to a medical provider; this approach generally involves some financial risk on the provider's part.) Contrasts in modes of reimbursement arise in part for historical reasons: since IPAs were formed to defend the fee-for-service sector, they were unlikely to pay physicians on any other basis. They also derive from an organizational logic. It is dangerous to capitate individual physicians, since this makes the incentive to underprovide care very stark. It is awkward to pool the risks of solo practitioners who do not work together or even know each other. And it is impractical to salary physicians who work for the HMO on an occasional and individually variable basis.

PGP physicians thus depend rather heavily on the HMO, which generates a large proportion of the patients they see. PGP physicians also interact on a continuous basis with other HMO physicians, who are their colleagues within the group practice or hospital staff. Financial rewards in PGP-IPAs are directly tied to HMO outcomes and policy through the setting of salary and capitation levels. (For capitated physicians, income is also directly tied to the practice behavior of other physicians, since groups of doctors generally form a risk pool.) And PGPs are able to influence physicians in a variety of other ways, including informal norm setting within the medical group, advocacy of practice patterns by the medical director, and formal attempts to develop coordinated practice patterns.

IPAs diverge from PGPs along all these dimensions. Their physicians have a weak financial and professional stake in the HMO, interact little with each other, receive modest financial incentives to connect their outcomes with those of the organization as a whole, and are relatively difficult to influence in informal or continuous ways. Management in these plans is frequently contracted out to an HMO management firm, rather than provided by organizational sponsors. Overall, IPAs can be so structurally disarticulated that they resemble a network of contracts more than an authoritatively coordinated organization.

A main result of these structural differences is that PGPs have historically succeeded in lowering health costs, while costs in IPAs are comparable to costs in the fee-for-service sector (Luft 1981). IPAs make frequent use of formal constraints on physician decision making like hospital admission requirements, but do not employ the combination of continuous and informal control mechanisms that translates into successful cost containment (Hillman, Pauly, and Kerstein 1989; Strong and Curran 1992). This situation may be changing, however, as is discussed in this chapter.

POPULATION DYNAMICS AMONG TYPES OF HMOs

It is of great interest that the IPA blossomed in the 1970s and 1980s. In the early 1970s, a handful of small IPAs was dwarfed by a larger number of well-established, high-enrollment prepaid group practices. But rates of IPA formation exceeded those of PGP formation in the 1970s and 1980s, with virtually all of the mid-1980s boom occurring through new IPA starts. IPA density rose to over 400 organizations in 1987, or two thirds of total HMO density. And IPA enrollment grew to virtually match PGP enrollment. HMO population counts by plan type are given in Figure 8-2.

One source of this shift in relative density is a challenge-response process. As already noted, the initial spur to IPA development was the threat posed by growing PGPs to the fee-for-service sector. This threat grew much more real with federal support for HMOs in the 1970s. IPA founding was thus particularly rapid in these metropolitan areas where PGPs had already appeared (Strang and Uden-Holman 1990). But this competitively induced effect declined in the 1980s (Wholey, Christianson, and Sanchez 1993) as HMOs became a regular and less threatening part of the medical scene. Although a challenge-response process can explain the initial takeoff in IPAs, it cannot account for their increasing predominance within the HMO population.

A second main source of the growth of the IPA subpopulation is its ease of founding and operation. IPAs rely heavily on the existing "building blocks" provided by the health system; most importantly, the ability of physicians to practice individually and through referrals in the absence of a central coordinating mechanism. IPAs are trivial to set up because they involve no necessary
change in physician routines, occasion no strong dependence on the plan for revenues (since physicians maintain their existing practices), and require little managerial expertise or structure. Of course, these are the attributes that led health policy analysts and reformers to disparage IPAs. But they are also attributes that fueled rapid construction once public health policy and market conditions began to favor HMOs.

On the other hand, prepaid group practices are difficult and costly to found. They involve substantial start-up costs, large commitments by physicians and others that can only be met as the plan grows, and specialized forms of managerial expertise not found in the health sector at large. PGP s are difficult to found and to maintain because they differ so importantly from conventional medical arrangements. Brown (1983) is forceful in arguing that the absence of rapid industry expansion in the 1970s was not due to faults in federal legislation, but to the very real costs and difficulties of constructing alternative kinds of health-care organizations.

By this argument, IPAs are what ecologists have dubbed R-strategists: organizations that are able to reproduce rapidly in newly accessible environments. (Their ease of (re)production is contingent on the prior existence of an operating fee-for-service system, of course.) PGP s are K-strategists: organizations with high fitness in stable or crowded environments, but with less capacity to rapidly exploit new opportunities. Some evidence for such an analysis is suggested by fact that IPA formation took off when capital markets made large influxes of money available to plans that could be founded in a hurry.

However, a third account suggests that IPAs have expanded rapidly due to improved operating procedures—that the modern IPA has simply built a better mousetrap. Welch (1987) argues persuasively that in the 1980s IPAs learned from well-publicized failures like that of SAFECO’s United Healthcare. Many began to provide a substantial fraction of the patients seen by their physicians, developed risk-sharing reimbursement practices, and instituted more thorough organizational controls.

Welch contends that contemporary IPAs are able to outcompete PGP s because they have come to combine the traditional strengths of the PGP with their structural advantage in flexibility and ease of start-up and expansion. While IPAs grew aggressively at the expense of established staff and group HMOs in some metropolitan markets, however, it is not clear how Welch’s arguments can be squared with the fact that IPA failure rates exceeded PGP failure rates throughout the 1980s.

A second set of interform dynamics involves transitions between staff, group, network, and IPA models. As one might expect from the previous discussion, there is little movement between the IPA form and the other three forms, but some movement between different PGP types. In particular, staff and group HMOs fairly often turn into network HMOs. This shift is largely a function of organizational expansion. For example, the growth of a network HMO leads to the development of contracts with additional group practices as much or more than through expansion in the size of the original group.

These last points serve as a useful caution about the permanence of organizational typologies. As Welch argues, aggressive IPAs have taken on many of the attributes traditionally connected to PGP s; in turn, the now less embattled PGP s may likely take on traits of IPAs (see next section in chapter for the example of Group Health Plan in the Twin Cities). As a result, knowledge of plan type conveys less about core internal processes within an HMO now than it once did. Standard industry classifications tend to become institutionalized. In the case of HMOs, the present four-type classification is grounded in federal policy and in Interstudy’s (1973–75) HMO census categories.

Welch, Himmel, and Pauly (1990) have argued that new HMO typologies should be developed that better reflect contemporary variations in organizational structure. They propose that the crucial dimensions of HMO structure include how physicians are reimbursed, whether physicians retain a practice outside the HMO, whether physicians contract with the HMO as a group or as individuals, and the type, size, and structure of risk-sharing arrangements between the plan and providers. This reconceptualization may be seen as an effort to more directly characterize the organizational structures of different types of HMOs, and in particular to distinguish between “new” and “old” IPAs.

**HMO Dynamics in One Metropolitan Market**

HMOs are involved in complex local systems of competition, imitation, and learning. The initial entry of an HMO into a community looks much like a random event, though it is conditioned by the community’s existing health sys-
HEALTH MAINTENANCE ORGANIZATIONS

much of this market. Total HMO enrollment in 1973 stood at about 65,000 persons, almost three quarters of whom were covered by Group Health. HMO enrollment in the Twin Cities rose to 275,000 in 1979, and stood at some one million subscribers in 1987. At that time, Group Health enrolled 228,000 persons, or about a quarter of all HMO enrollees. The largest plans were now MedCenters (now linked to the PARTNERS national network), with 292,000 enrollees, and Physicians Health Plan, with 324,000 enrollees.

Active competition among HMOs produced a number of important shifts in the internal operations and market offerings of the Twin Cities plans. First, competitive pressure on health costs lowered physician incomes, leading to a "physicians' rebellion" (Kralewski et al. 1987). When PHP failed to return any funds from its physicians' withholding account in 1987, its doctors mounted a campaign to oust the plan's board of directors that succeeded in gaining partial control of the organization. Competitively induced decreases in MedCenter's premiums led to an arbitration fight over physician incomes. HMO Minnesota failed in the late 1980s as a result of rising competition.

While IPAs faced mounting internal tensions in an increasingly tight market, the older, highly integrated plans faced pressure as well. As noted in the figures on enrollment, Group Health Plan was increasingly superseded by the more flexible group and IPA plans. Aggressive IPAs like PHP offered similar premiums and much greater choice of physician and hospital than even a large-staff HMO could make available. As a result, in 1987 Group Health began to offer preferred provider options, where enrollees could see physicians outside the plan (but pay higher rates for such visits).

Finally, Minnesota's Blue Cross responded to HMO growth by adopting many of the practices of HMOs. In the late 1980s, Blue Cross began to aggressively contract with hospitals rather than paying standard fees, and to withhold 10 percent of physicians' income against plan losses. As a result, Minneapolis-St. Paul's conventional fee-for-service sector has moved to a form of "managed care" that is hard to distinguish from the structure of many IPAs.

This brief account of HMOs in the Twin Cities illustrates several population dynamics of broad import. Entry of HMOs into the metropolitan market was slow and contested until the 1970s, when new federal policies and trends in the medical market suggested that more organized forms of health delivery were imminent. A variety of group practice-based, and then solo practitioner-based, HMOs were formed. Most have prospered, but mounting competition leaves decreasing slack for all types of health organizations. The conclusion of some seasoned observers of health care in Minneapolis-St. Paul is that tight competition homogenizes the internal structure and market products of health-care organizations (Fieldman, Kralewski, and Dowd 1986).

CONCLUSION

Scott (1987) notes that organizations may face complex technical environments and complex institutional environments, and illustrates the point with the example of health care. Due to their efforts to coordinate and organize aspects of
health care usually left to the market or informal relations between physicians. HMOs face more extreme environmental pressures than other medical organizations. Health maintenance organizations must manage a complex and uncertain work process, enroll members in an increasingly tight medical market, and locate their operations within a highly politicized, wealthy industry viewed widely as in crisis.

While the care delivered by HMOs is strongly conditioned by the changing technology of medicine, technological change appears to have played almost no role in the population dynamics of HMOs. Nationwide, it is true that the increasing specialization and complexity of health care has led to the expansion of group practices and the demise of the independent solo practitioner. But this has not had a discernible impact on the HMO industry. As discussed, the largest increase in health maintenance organizations has been through the organization of contractual arrangements among the solo practitioners forming IPAs.

Instead, the historical dynamics of HMOs are most importantly a function of institutional conditions. Prior to 1970, HMO growth was stunted by the political power of the medical profession. Since 1970, health maintenance organizations have received strong support from government policy, not in proportion to their political strength but in proportion to their attractiveness as vehicles for cost containment. Public policy support generated unexpected outcomes, as types of HMOs regarded as little able to contain costs proved the best capable of seizing new opportunities. Given the volatile character of national health policy, institutionally generated change seems likely to continue to dominate the population dynamics of health maintenance organizations.

NOTES

1. A detailed historical portrait of Kaiser, from which this discussion is drawn, is provided by Smillie (1991).

2. Incentives were particularly weak because most United Healthcare doctors had "panels" of less than twenty UHC patients. Whether physicians received money back was primarily dictated by whether they happened to see one or two really ill patients. Further, it was difficult to spread physician risks because UHC physicians were primarily solo practitioners who were unable to exert informal controls over each other, increasing the likelihood that physicians would free-ride on any collective risk pool.

3. In fact, a 1980 survey showed that physicians in IPA talk less frequently with each other than they do with physicians outside the HMO.

BIBLIOGRAPHY


Newspaper Publishers

GLEN R. CARROLL

Newspapers have been part of American life for over three hundred years. The beginning of the industry is usually placed at 1690, which marked the appearance of the first known American paper, Publick Occurrences Both Foreign and Domestick. Since then, thousands of newspapers have been started in the United States, penetrating virtually every locale. The industry’s period of greatest organizational growth was the nineteenth century, when the number of papers went from 150 to over 20,000. Most counts suggest that the number of newspapers in the country peaked at about 23,000 immediately prior to World War I. Throughout the rest of the twentieth century, numbers of American papers have generally declined.

The last decade or so has witnessed a major development in the newspaper industry—the emergence of a national press in the United States. Although the Christian Science Monitor and the Wall Street Journal had long operated on a national basis, they were considered specialized anomalies. The launching of Gannett’s USA Today in 1982 and the subsequent expansion of the New York Times into many far-flung metropolitan areas represented a shift in size and scale of potential newspaper markets. According to many observers, this shift was made possible by modern advancements in information technology.

Prior to the national press’s emergence, newspapers in the United States typically focused on a single metropolitan area. This is still true of most American newspapers. However, the persisting local character of the press does not mean that the organizational structure of the industry has remained stable. Indeed, the transformations of the industry across the many and varied places of the country are remarkably similar and dramatic. They also foreshadow the national press, making it appear as the natural outgrowth of a long-term trend.

This chapter reviews organizational transformations that occurred in the American local newspaper industry in the years prior to the development of the national press (up to about 1975). It describes several long-term trends in