The Rise and Decline of the Early HMOs: Some International Experiences

BRIAN ABEL-SMITH

London School of Economics and Political Science

HEALTH MAINTENANCE ORGANIZATIONS (HMOs) have gradually been taking an increasing share of the health insurance market in the United States, though they are still far from having the largest share. They are seen by many commentators as a relatively new development. Many have argued that they offer a long-run solution to the problems of the medical market by generating greater "efficiency" and controlling costs. Thus, they represent a "best buy" among private health insurers. Their claimed advantages are such that some authors in the United States are advocating them as *the* solution to health financing which should be adopted in other countries—even in developing countries (Akin et al. 1987, 33). There has also been substantial interest among academic authors in Europe and some isolated experiments.

The term "health maintenance organization" was coined around 1970 and legislation was enacted to promote their development in 1973. Without pursuing the complex subtleties of definition, it can be said that they combine two quite different categories of insurer, classified by the International Labor Office (ILO) many years ago as "direct" and "indirect." The direct insurer employs its health professionals and the indirect insurer contracts self-employed health professionals. This distinction is generally made in the United States less succinctly by distinguishing between prepaid group practices (PGPs),

The Milbank Quarterly, Vol. 66, No. 4, 1988

which employ doctors, or group model PGPs, which contract them, and preferred provider organizations (PPOs) or independent practice associations (IPAs), which have more varied types of contracts (Brown 1983). In the former, doctors are selected, trained, and employed to be low authorizers of health services. They are paid by salary with or without a bonus. In the second, the contracted doctors are also expected to be low authorizers of health services; if they are not, their contracts will not be renewed. They are paid by salary (with or without a bonus), capitation or, in the case of PPOs and IPAs, by fee-for-service generally somewhat below prevailing levels of fees.

As will be shown, in the period between the two world wars, direct and indirect insurance, both of which combined the functions of financing and securing the provision of health care, were by far the most common forms of health insurance in the world. The third ILO model, reimbursement of the insured, was by far the less common model. As the history of health insurance has been different in the United States, the reimbursement model by a third party is often referred to as "conventional health insurance" (Brown 1983). In the period between the two world wars, HMOs were regarded as the conventional form of health insurance outside the United States. If they are now believed to provide the solution to securing quality health care at acceptable cost, why are so few competitive HMOs now to be found outside the United States? While the development of HMOs in the United States has been a "long, precarious, risky, business" (Brown 1983), this was not the case in Europe. They not only dominated the insurance market but grew to cover one-third to one-half the population in some countries by 1933. Why did they expand so readily without great risk? And if they are the best solution to the problems of cost and quality, why did they not continue to dominate the European market for health insurance? Why did their role decline? Or why was it terminated? These questions are pursued in the case of those European countries with a high coverage of health insurance by 1933.

Before investigating these questions, it is necessary to spell out in greater detail the characteristics of HMOs to see how far the international comparison is really valid. First, HMOs operate in a competitive environment; they compete for patronage. Second, they generally adopt some form of community rating, though the precise meaning of this term may vary. For example, they may restrict their

operations geographically to avoid areas where the poor live. Third, they operate on a closed-panel basis. No local doctor has a right to join the plan. Fourth, they do not normally pay doctors fee-for-service. Fifth, they generally operate on a referral basis. Specialist and hospital care is only paid for if authorized by the closed-panel doctor originally consulted. This is the key to managed care—to the economies secured mainly, if not exclusively, in hospital use. A sixth possible feature implied by the title is an accent on prevention, but this is not borne out in practice to any great extent.

The conditions which are necessary for HMOs to offer their competitive edge are the following. First, consumers must have a free choice of insurer. Second, the organizations must be in a position to hire or contract selected doctors and dismiss them from their association with the organization if they do not perform as intended. The organizations have, however, to maintain the goodwill of their subscribers or they will not rejoin next year.

The Rise of the European HMOs

Historical Development

Health insurance covered over one-quarter of the population of Switzerland by 1903 (Hogarth 1963), over 20 percent of the population of Great Britain by 1910 (Green 1985), 35 percent of the population of Denmark by 1911, 18 percent of the population of Germany by 1914, and 20 percent of the population of Norway by 1920 (Hogarth 1963). In only two of these countries (Germany and Norway) was health insurance compulsory in the years indicated.

Broadly comparable figures for the year 1933 are available. They are shown in table 1.

In four of these countries (Denmark, the Netherlands, Sweden, and Switzerland) virtually all health insurance was still voluntary in 1933.

The early development of health insurance in many European countries was in association with the guilds. Thus, one health insurance organization was generally established by each guild for its members. This precedent of basing health insurance on occupation was continued in some countries such as Germany, Austria, and partially in Switzerland and Sweden. It was encouraged in Germany by legislation requiring particular types of employer to develop and contribute to

TABLE 1
Coverage of Health Insurance[2] (1933)

	Population (in millions)	Insured (in millions)	Insured as % of population
United Kingdom	46.0	17.7	26%
Germany	65.0	18.5	28
France	41.0	16.0	39
Netherlands[1]	8.6	3.5	41
Denmark	3.5	2.4	68
Switzerland	4.0	1.8	45

[1] Source: CBS 1937 in: van der Hoeven, H.C., 1983. *Om de Macht bij het Fonds* 168, Den Haag: Azivo.
[2] Source for countries other than The Netherlands: Levy, H., 1944. *National Health Insurance.* Cambridge: Cambridge University Press.

health insurance for their employees. The classic early case dating from the eighteenth century was that relating to German mine owners. This pioneering development was followed by other employers, and the occupational basis of insurance funds was maintained when a compulsory scheme of health insurance was introduced in Germany in 1883. The gaps in occupational funds were filled by new specially created insurers covering geographical areas. Thus, from the start employees were, with some exceptions, denied free choice of insurer so that competition for custom was restricted.

In northern Europe after the guild system fell into disuse or was abolished, it was small farmers and working men themselves who created their health insurance agencies. This was the case in Norway, Sweden, the Netherlands, the United Kingdom, and to some extent in Switzerland. When a law was passed in Denmark in 1861 to abolish the guild system, it was the medical profession which lobbied to preserve the mutual aid societies associated with them. The profession also played a leading role in establishing the early funds and thus had inside control of their policies (Rasmussen 1979, 48). In France the mutual insurance agencies were more middle class and not supported by the trade union movement (Saint-Jours 1983). Health insurance in England was the initiative of local groups of workers—originally those with regular work in the more skilled occupations. In time some grew to be large national organizations with local branches. Hospitals did not occupy the important position in the care of patients which they do today. Moreover, when general hospital

care or specialist consultation was needed it was generally available through charitable hospitals originally provided free and later provided on an informal means-tested basis. Thus, hospital and specialist care was very rarely covered by insurance. The largest category of health insurers involved "friendly societies" which provided cash benefits for the sick, treatment from contracted general practitioners, and drugs. Later, some dental and ophthalmic benefits were added by some societies. Other insurers included trade union clubs, slate clubs or tontines (clubs which distributed any annual surplus among the members), works clubs (based on the factory), and provident dispensaries which were subsidized by charitable funds. There were also some medical aid societies run on similar lines by commercial life insurance companies (Green 1985), but they were in the minority. Membership grew to include not only the mass of wage earners but small shopkeepers and a substantial proportion of the middle classes. The key point was that any reasonably healthy male worker could join any club provided he was willing to pay the contribution. And once a member he could stay a member even if his health deteriorated seriously. Some clubs admitted working women, wives, and children. There were thousands of different clubs competing for members.

Dutch medieval guilds were abolished at the end of the eighteenth century. Later, private commercial insurers tried to provide health insurance but experience with these was unsatisfactory (Hogarth 1963) and in the second half of the nineteenth century voluntary health insurance run by insured persons themselves or sometimes by industrial firms or philanthropic bodies was established. By the end of the century there were hundreds of such funds of varying size, structure, and resources competing for members. Benefits included hospital care. The growth of health insurance at the initiative of farmers and workers as well as employers was also a response to the bad reputation of private insurers in Switzerland. Insurance agents were a sort of peddler who went from house to house seeking orders, and they were paid considerable commissions. "Often people with unsuitable characters were employed who promised all sorts of things before concluding a policy which were not realised when a claim was made" (Maurer 1983). Health insurance included hospital care often provided by subsidized cantonal hospitals. Voluntary health insurance developed just after the middle of the nineteenth century in both Denmark and Sweden and somewhat later in Norway (Hogarth 1963). In Sweden insurance was often based on a particular trade or occupation. By

1890 there were about 1,500 different funds in Denmark. From 1892 the funds were offered state subsidies based on membership and subscription income, heavily subsidized treatment for members at public hospitals, and free ambulance services (League of Nations 1924, 182).

Combined Finance and Organization

The easiest way for a commercial insurer to operate is to reimburse bills in whole or in part after they have been paid, just as it pays to replace property after a fire. But the consumer-run insurers did not want just to spread risks but to obtain cover on favorable terms. Thus, much or all of the administrative work of the early small insurers was originally done by the members themselves to cut costs and retain links with other members. In some cases—as in England, Norway, and Sweden—the original aim was to provide a cash benefit when earnings were interrupted by sickness. At a later stage a doctor was brought in to certify sickness and to shorten the duration of benefit by his treatment. Thus, selected local doctors were contracted to provide this service. The origin of capitation payment was not so much to contain costs as to simplify administration and make the cost of the medical benefit predictable so that a share known in advance of each member's contribution could be set aside as payment for the doctor. In the United Kingdom the doctor was originally expected to provide in return for this payment not only his services but his medicines. Capitation payment was originally used by the vast majority of insurers in Germany, Denmark, the Netherlands, and the United Kingdom. Thus, the functions of finance and organizing services was combined in one agency. Reimbursement came to be used only for a minority of the higher income groups in Denmark and Switzerland and much later in the United Kingdom. In France, as explained later, reimbursement was used when health insurance became compulsory.

Community Rating and Open Enrollment

A further advantage of capitation payment was that it made it possible, at least for the medical benefit, to combine community rating with open enrollment—one of the central problems facing those seeking to promote the development of HMOs in the United States (Brown

1983). Admitting and retaining members with a high risk of needing health care placed extra burdens on the doctor. But the doctor was, as we shall see, in a weak position to try to negotiate the level of capitation payment on this account. The health insurers did generally exclude applicants for membership who were sick at the time of application, as it was obviously unwise to admit a member who would immediately claim a cash benefit for sickness. One way of overcoming the problem was to lay it down that the benefit could not be drawn until the member had contributed for a specified number of weeks or months.

In 1892 the Danish government recognized the problem of persons already sick who could not get insurance and gave subsidies for voluntary insurers to accept them—the act laid it down that all persons had the right to insure themselves against sickness; the compulsory benefits were free medical and hospital care and a cash sickness allowance (Hogarth 1963).

Referral and Closed Panels

In the early European HMOs, capitation payment of the doctor imposed referral. An insured person had the right of direct access only to the doctor with whom he was registered. In a small society this may have been only one doctor or one in a particular area. When there were several doctors, an insured person had to make a choice among a closed panel, though limits might be placed on the number of insured persons a particular doctor could accept. Moreover, a period of notice was required for a member wishing to change doctors, or changes were only allowed periodically—for example, once a year. This stopped the patient from shopping around among different doctors during the course of an illness, which could happen under fee-for-service payment. It should also be mentioned that referral had become the medical convention in many north European countries where the specialist was largely or wholly hospital based. Capitation payment reenforced this convention.

Managed Care

Did the early HMOs manage care? In the early model in the United Kingdom when health care consisted only of access to the doctor and

the medicines that the doctor provided, management consisted in negotiating a low capitation payment for this inclusive service. When medicines became separately provided through pharmacists, the insurer had a clear interest in monitoring the prescriptions dispensed at the pharmacy and comparing the performance of participating doctors in this respect. Alternatively, the doctor was given what was left in the fund after the bills for medicines were paid. Utilization reviews and other controls were gradually developed in other countries where hospital and specialist fees were paid. Consumer-run organizations inevitably sought strong justification for authorizing an increase in the premium.

Provision of Services

Were both the direct and indirect method of securing services used? The most common model was indirect: doctors were contracted on a capitation basis (fee-for-service in Switzerland) and continued with their private practice among the better off. The advantage was that the insurer was not responsible for the total remuneration of the doctor and thus could secure his services on favorable terms, leaving him to supplement this by private-practice earnings which yielded much more per hour of work. The disadvantage was in the perceived quality of service obtained. There was always a risk that the insured would be given little time and low-quality medicines. It was largely for this reason that the direct method was also used in a minority of HMOs. These were certainly to be found in the United Kingdom. Some work funds maintained their own dispensaries, gave their doctors housing and fixed salaries, and built and managed their own hospitals (*British Medical Journal* 1905). But it was in the Netherlands that they played a considerable, if minority, role, not only hiring doctors on salaries but maintaining their own pharmacies, dental clinics, optician services, maternity services, and hospitals. In this "staff" model the insurer owned the premises and hired professionals to work in them.

How Did the Early HMOs Obtain Such a Large Share of the Market?

In summary, the early consumer-controlled health insurers in northern Europe had all the major characteristics of HMOs. They combined

financing and delivery. They were based on closed panels and a system of referral. They managed care. The financial incentives not to overserve were severe. Any time devoted to contract practice was time not available for ordinary private practice. As the doctor had originally to provide the drugs, every penny of unnecessary medication came out of his own pocket. Even where these were, in part, separately accounted for (e.g., expensive drugs), the total cost had to be contained within the allocation for the medical benefit. The doctor who failed to give satisfaction in balancing service against cost could find his contract terminated. There were both staff models (if a minority as in the United States) and the much more common system of contracting doctors working in their own offices. And they competed with each other for patronage. It is true that they did not generally provide preventive services except for some vaccination. But apart from this, little personal preventive care existed to be offered. It was before the age of prenatal care, child health monitoring, or family planning. Most important, they combined community rating and open enrollment to an extent not now thought possible in the United States.

Why were they able to procure a larger share of population coverage than HMOs have so far achieved in the United States? The early European HMOs had a number of advantages. Private commercial insurers were not in this market or only to a small extent. They had tried to enter it in both the Netherlands and Switzerland and became discredited. They did enter the market on a greater scale under National Health Insurance in the United Kingdom (1911–1948) but had, under the legislation, to have separate sections nominally subject to the control of their members and to operate on a nonprofit basis. Second, the persons covered did not generally have previous regular relationships with their own doctors. Indeed, joining health insurance made this economically possible for the first time. Third, there was no shortage of doctors, though some may have been poorly qualified, and it was not easy for some to make a living. Fourth, the technology of medicine was much less developed, which simplified the task facing insurers. Finally, the medical profession was as yet poorly organized. Indeed, as we shall see, it was the issue of HMOs more than any other which galvanized doctors into building up their associations into powerful organizations equipped to fight for their interests. Their

objective was, in one way or another, to stop insurers being in competition and to take control of the way in which they operated.

The Battle against Competition

The battle against competition was fought by doctors throughout Europe in the twentieth century. Indeed, it was known in Britain as "The Battle of the Clubs". And the agency which the doctors eventually persuaded to rescue them from market forces was government. It is possible that they could have won without it, as is shown later from the experience of Australia. But the governments of Europe did not stay inactive long enough to see whether this would happen. It was because governments were anxious to extend the coverage of health insurance and maintain its continuity in one way or another that they intervened in the medical market and responded to pressures from the doctors. As we shall show, the abolition of competition was the price governments found themselves paying to secure wider coverage.

Competition hurt doctors where it really mattered—in their incomes and the workload needed to maintain them. While some insurers contracted doctors at rates regarded as fair by both parties, in other cases the job of contract doctor was put out to tender and often, though not always, awarded to the lowest bidder. Well-established local doctors felt obliged to tender lest a rival come into the district who could start up his practice on the minimal living offered by insurance practice and hope to add gradually to it by attracting private-practice customers at the expense of long-established practitioners. Taking the contract involved the doctor in no extra overhead costs. Thus, he could afford to offer marginal cost pricing. Doctors usually did not have the security of tenure. They typically held office at the pleasure of the appointing board. They could readily be dismissed even for the reason that other doctors were prepared to take the contract at a lower price. As a market mechanism for obtaining medical care at the lowest cost, the system could hardly be improved upon.

In the United Kingdom it was not only health insurers who used the system of contracted doctors. It was also used by the local poor-law authorities who put out to tender the part-time salaried jobs as district medical officers whose duties were to provide medical services to whatever number of persons had been designated as indigent and

needed what was called "medical relief." These medical officers also generally had to provide drugs at their own cost as part of the contract (Abel-Smith 1964). They also had no security of tenure. This remained the system in most poor-law districts of "unions" until as late as 1909 (Royal Commission on the Poor Laws and Relief of Distress 1909). Thus, both the private insurers and the poor-law authorities were operating as health maintenance organizations.

The opposition of the profession to the system of what was called "contract practice" is well documented in the United Kingdom. The profession objected to being subject to "impudent" patients demanding drugs and "impertinent" lay committees who decided on the renewal of contracts (*British Medical Journal* 1905). Second, the profession wanted to establish a national capitation fee of five shillings per member per year compared with sums as low as two to three shillings or even less for which some doctors had agreed to take such contracts (Green 1985). Third, they objected to the further loss of income when persons who could well afford private practice joined or remained members of the clubs. The "friendly societies" were, however, unwilling to allow the profession to rewrite the rules of their societies. They were particularly unwilling to apply income limits for membership, particularly when many of the higher-income members were long-established members whose economic circumstances had improved. Moreover, some of them held positions of leadership in the societies or "clubs" (Green 1985). Fourth, the interests of younger doctors denied access to patients because of the closed panels were represented. The profession wanted the right of any doctor to take insurance patients.

The issues raised in the Netherlands were similar. There, the medical association forbade its members from accepting contracts under a specified annual sum per patient. But inevitably some doctors did so out of economic necessity. It proved impossible to secure standardization of payments at a level which satisfied the profession. Second, the profession also wanted to establish income levels for membership. Third, the medical association wanted at least one-half the seats on the boards of insurers to be allocated to members of the medical or pharmaceutical profession. It was argued in 1913 that "a sickness fund dominated by the insured will tend to maximize medical services for minimal charges" (van der Hoeven 1983). Laymen would favor a narrow definition of medical services out of ignorance and a

desire to keep costs low. Fourth, the profession wanted to separate the provision of health services from cash benefit to the sick. Where cash benefit was included, the triangle of trade unions, employers, and the state was believed likely to dominate the control of the organizations (Juffermans 1982, 124, 154). In Denmark also, the main issues raised by the profession were the level of capitation payment, conditions of employment, and the protection of clinical freedom (Hanson 1974, 121–22).

Rival Insurers

The first strategy adopted by the medical profession to try to protect its interests was to establish rival insurers under their own control. There is a parallel with this tactic from the experience of the United States. As Brown (1983) put it, "An IPA will follow an HMO in a community only a little less automatically than night follows day." In the United Kingdom, doctor-controlled insurers were given the misleading title of "public medical services." Usually they were founded by the local branch of the British Medical Association from 1895 onwards (*British Medical Journal* 1905). But they failed to secure more than a small share of the market. In the Netherlands the initiative came a little later and was more successful. There, the Medical Association started a central organization to establish doctor-controlled funds in 1913. They were openly called "Association Funds" and the aim was to establish one in every district (van der Hoeven 1983). Progress was slow at first, but in 1937 there were 84 funds with 1,180,000 members—nearly one-third of the market. The country where the medical profession did, if much later, manage to secure the largest share of the market for medical insurance was Canada, where 4.8 million persons were covered by 1961 (Royal Commission on Health Services 1964). But this did not succeed in stopping the government from introducing publicly controlled compulsory health insurance.

Control through Licensure

A second strategy in the United Kingdom was the attempt to persuade the national licensing authority (the General Medical Council [GMC]) to intervene and declare service with a medical institute "infamous

conduct," thus leading to removal from the medical register. But the most the "Council" could be induced to do was to express in 1899 its "strong disapproval" of medical practitioners who associated with those medical aid associations which systematically canvassed and advertised with the purpose of procuring patients (Green 1985). By 1902 both canvassing and advertising were, however, declared "infamous conduct" (Green 1985). This was a serious blow to the provision of the information needed for consumers to exercise free choice. Members of the GMC argued, on the other hand, that they had no power to protect doctors' "pecuniary interests" (Green 1985).

Boycott

The profession then had to fall back on the third strategy of trying to boycott insurers offering jobs on terms to which the profession objected and refusing to accept advertisements for such appointments in medical journals. This tactic was only occasionally successful in some districts because doctors were often brought in from outside who had not agreed to participate in the boycott. In the Netherlands, the medical association fought a continuing battle to boycott the health insurers it did not control. A particular target were the 68 insurance funds associated with the trade union movement with over half a million members which worked together as a federation. They were singled out for boycott because many of them used the staff model—hiring doctors and other professionals on a salaried basis. This was seen as an attack on the financial and professional independence of the professions (van der Hoeven 1983). The battle was particularly fierce in the early 1930s. Whether this boycott strategy would have eventually been successful in the United Kingdom or the Netherlands will never be known as in both countries the introduction of compulsory health insurance ended the war, as described below.

What is however interesting is that the strategy in a country outside Europe—Australia—did succeed. Voluntary health insurance covered 6 percent of the population of that country by 1900. Consumer-run insurers (friendly societies) had developed from the 1830s providing capitation-paid medical services and drugs. Their development had originally been welcomed by the doctors but by 1900 they criticized them on similar grounds as elsewhere. First, the level of capitation payment paid by many societies was considered to be far too low.

Second, doctors objected to being subject to control by lay committees. Third, they resented high-income earners and their families being admitted as members. In 1909 the New South Wales branch of the medical association managed to impose a common form of agreement on the local societies with a stated level of capitation payment, the right to charge extra for operations, anaesthetics, and confinements, and income limits for membership (Sax 1984). In 1918, 406 doctors in Victoria resigned from the Societies. They refused to rejoin until the Societies finally accepted, in 1920, a common form of agreement similar to that of New South Wales. By 1922 all state branches of the Association had negotiated such agreements (Sax 1984). The result was a rise in subscriptions and a fall in membership. By boycott the Australian doctors managed to achieve the same objectives which the British doctors had been seeking and did not obtain until the government made insurance compulsory.

Soon afterwards the profession enunciated the principle that the provision of medical benefit should be in the form of cash rather than kind: patients should pay fees to doctors so that they would have a sole loyalty to their patients, not to any third party such as the government or a friendly society. Patients must also have a free choice of doctor. In 1949 the medical association terminated all contracts with friendly societies (Sax 1984). Not until 1953 did they agree to make new contracts, by which time the societies had conceded the new basis of contract which the profession was demanding.

Government Regulation

The fourth way in which doctors could protect themselves from the consequences of a competitive market was by securing some form of government regulation. The problems of the Danish doctors were largely resolved by a law enacted in 1915 which established an arbitration tribunal to deal with disputes between doctors and insurers. Arbitration inevitably led to standard capitation payments—the removal of price competition. The same act dealt with a further grievance of the profession. Separate unsubsidized and state-regulated insurers were established to provide for those with incomes above a specified level. They were called "continuation funds" (Hogarth 1963). The consequential "wastefulness" of having well over a thousand funds providing similar benefits on the same terms was eventually recog-

nized. An act of 1960 secured that only one fund could exist in each commune to save administrative costs, with the exception of certain funds covering the whole country. This act recognized that by this time competition had ceased to have any function.

Similarly, the Swiss doctors managed to secure legislation to protect their interests at about the same time. A federal law to establish compulsory sickness and accident insurance had been passed in 1899 but was overwhelmingly rejected by a plebiscite in 1900 (Maurer 1983). But a law tightly regulating voluntary health insurance was passed in 1912. It conceded the main demands of the doctors. Under it, patients were to have free choice of doctor and doctors were to be paid on a standard tariff laid down by each local government (canton) after consultation with the interested parties, either of which could take the question of fee levels to arbitration. To attract a federal subsidy, a health insurer had to meet certain conditions—provide minimum benefits, calculate premiums on the age when the person first entered insurance, not charge women premiums more than 10 percent above those charged for men, and allow all persons (except those with preexisting health problems) to join irrespective of health status. Moreover, a member who changed residence or occupation has to be accepted by a fund in a new area without extra contributions based on his health or age.

This act laid the basis on which voluntary health insurance still operates today. While there are funds open to anyone, some are confined to certain employers or occupations (Abel-Smith 1985). Swiss health insurers do not, however, operate as independent competing HMOs except in the range of extra benefits they offer. But in each canton there is an association of health funds which negotiates price agreements with cantonal doctors associations and with cantonal hospital carriers; some other contracts are negotiated at the national level (Abel-Smith 1985). Persons who wish to do so can pay extra premiums for single rooms in hospitals and access to doctors who charge higher fees. But these higher fees are also negotiated (Abel-Smith 1985). From 1912 onwards medical benefit was made a standard cantonwide package. This largely overcame the problem of risk rating but it barred effective price competition.

German doctors faced similar problems to their colleagues in northern Europe, even though health insurance had been made compulsory for wage earners in the larger industries by Bismarck in 1883 and for

further employees later on. The question of an income limit was, however, resolved as part of the original legislation and an income limit has been retained ever since. While German sick funds were not generally in competition for members as the main occupations had their own fund with the residue joining one local fund, they did originally act like HMOs in using closed panels and trying to buy quality care at low cost. Benefits included treatment from contracted doctors and hospital care and the persons covered were gradually extended to cover family members, normally without any extra contribution (Zollner 1983). Originally there were about 20,000 different funds and each could nominate particular doctors to give services. Doctors started to undercut each other for contracts and the funds took advantage of this situation (Zollner 1983). But the doctors campaigned from 1900 onwards to stop this practice and, after a threatened national strike of doctors, from 1913 the entry of doctors to health insurance became controlled by joint committees representing the doctors and the funds (Hogarth 1963). Moreover, any contract with a doctor required the approval of a contract board consisting of an equal number of representatives of the doctors and the insurers (Zollner 1983). On the other hand, medical referees were employed by the sick funds to control expenditure on medical benefits (Hogarth 1963). From 1932 the profession obtained a decree which specified that no doctor had any contract or responsibility to any one fund. The various funds paid a negotiated level of remuneration on their membership to a professional body (the Insurance Doctors Association) established in each province which distributed this sum among doctors according to work done. From 1960 all doctors could participate (Zollner 1983). Thus, the profession established itself as a monopoly to provide treatment (Hogarth 1963) and evolved the schedule of fees used to pay for work done. Sick funds were prevented from acting as separate HMOs controlling their own payments for services. Not only was there little competition for members but no competition for the pricing of services.

The rescue of the British doctors came as part of the package which established national health insurance. When Lloyd George announced in 1910 the government's decision to introduce the scheme, the British Medical Association was galvanized into concerted action. It laid down the principles upon which it was prepared to collaborate with the

new proposed legislation. They presented their grievances and their proposed solutions to them which had accumulated over more than half a century. On no account would they be prepared to cooperate with legislation which involved no more than requiring employees to join any one of the existing organizations offering health insurance for prescribed benefits. In short, they threatened to boycott the scheme. What they wanted were income limits for persons entitled to medical benefits, free choice of doctor by the insured person, adequate remuneration, and the medical benefit to be administered by statutory committees on which doctors would be strongly represented. This last point was the critical one. No longer should friendly societies make the contracts with doctors or deal with questions of dismissal (Levy 1944).

In the course of negotiations the government eventually conceded all of these principles in whole or in part. Local statutory bodies ("insurance committees") were established to contract doctors on much more favorable terms negotiated centrally. The initial capitation payment was finally fixed at 7 shillings per insured person plus two shillings for drugs, which can be compared with the 5 shilling level which the BMA had been trying to impose on the friendly societies (Harris 1946, 215). An income limit for compulsory membership was established. Any doctor could contract patients. The only disappointment was that doctors only obtained a minority of places on the statutory "insurance committees." While the consumer had gained free choice of doctor, he had lost free choice of insurer. Moreover, the policies of each local monopoly insurer were in part determined by representatives of the doctors. The removal of a doctor from insurance practice could only be for serious offenses against the regulations. Thus, the British doctors obtained at the start what German doctors did not achieve until 1960—after nearly 80 years experience of compulsory health insurance.

The rescue of the Dutch doctors came much later but by the same route. It was under German occupation that compulsory health insurance was introduced in 1941. Income limits, the abolition of closed panels or freedom for each doctor to participate, and standardization of premiums and contracts was imposed. An equalization fund forced the cheaper funds to subsidize the more expensive (Juffermans 1982). The staff model was at first discouraged and later it became impossible

to establish new funds on this basis (van der Hoeven 1983). And it was only these few remaining funds which competed for members. The 1941 law almost put an end to competition.

In Norway voluntary health insurance developed later than in the other north European countries. It was also on a smaller scale and mainly provided cash benefits. Although the Norwegian Medical Association was also late to develop and initially weak, its founders concentrated their efforts on influencing legislation (Bjercke 1970). They managed to place their members on important legislative committees. When in 1909 plans were being made for compulsory health insurance, the profession secured representation on the committee preparing the legislation (Aubert 1970). They were thus able to secure free choice of doctor on the ground that closed panels would create a privileged class of doctor. They also secured that payment should be on a fee-for-service basis, that doctors should determine their own fees, which would be fully paid by the insurer, and that the doctor should decide what was the most appropriate treatment (Urdal 1961, 114–26). Third, they secured that there would be only one fund in each locality. As a result, there was no competition and doctors were free "to abuse the system established by the 1909 Act" (Berg 1986). Conflict between the profession and the funds continued year after year without reaching common ground until an economic crisis in the mid 1920s led to the introduction of cost sharing: the patient was reimbursed only up to an official fee scale.

A similar solution was adopted in France by boycotting the government's plan for compulsory health insurance. As in Norway, France had, as late as 1914, a much less developed system of health insurance than most of its eastern and northern neighbors. This was partly because it was not until 1898 that voluntary mutual insurers were allowed to operate outside government control. Moreover, the membership tended to be drawn from the middle class (Saint-Jours 1983).

The issue of compulsory health insurance was forced on the political agenda when the provinces of Alsace and Lorraine were returned to France as part of the settlement after World War I. Thus, part of what had become the French population were covered by German health insurance. A law for compulsory health insurance was passed in 1920 but it encountered such opposition that no scheme was implemented until 1930 (Saint-Jours 1983). Part of the delay was due to differences between employers and employees. The employers

wanted exclusive control of any scheme and so did the mainly middle-class mutual insurers. Neither of these solutions was acceptable to the trade unions (Saint-Jours 1983). But the scheme was also opposed by the medical profession. The original intention was to pay doctors on a capitation basis and the profession was insulted by the low level of payment originally suggested. The profession also thought payments under the German scheme were far too low. The final scheme which was introduced reimbursed patients for part of their doctors' bills. The argument was one of principle: no third party should stand between the patient and the doctor. But this key principle secured at a stroke two underlying objectives—free choice of doctor and, crucially in view of German experience, free entry for all doctors. Third, it led to local negotiations undertaken by representatives of insurers in the hope of securing agreement about the total fee, part of which was to be reimbursed. This had the effect of preventing separate insurers from trying to get doctors to undercut each other's fees, while leaving individual doctors free to charge more to wealthier patients or to all their patients. There was little incentive to under-charge and considerable temptation to overcharge. This issue of what became called in Canada "over-billing" by some doctors became a continuing source of strife between government and the profession. In France as in Britain the coming of compulsory health insurance destroyed the ability of the funds to act as competitive HMOs.

Early HMOs in the United States

Why did developments of this kind not occur in the United States? The answer is that they did, if only on a limited scale. In one way or another the medical profession, rather than try to persuade the first generation of HMOs to operate in ways they found acceptable or get government to force them to do so, managed by boycott to cut down their operations to a negligible size. Those who identify the first HMO as that in Elk City in the late 1920s overlook the first generation of HMOs established on European lines. As in Europe, by the turn of the century certain large companies, such as railways or mines (particularly those operating in isolated and unsettled areas) provided medical care for their employees. Some built their own hospitals and clinics but most arranged for treatment through independent doctors and hospitals for a flat rate per month—the European system of

capitation payment. Companies were able to get doctors "to bid against each other and price down the value of their labor" (Starr 1982). But somehow the opposition of the medical profession managed to prevent any major expansion of this type of service and eliminate or change beyond recognition earlier plans which provided it. In at least one case a doctor working in this type of service was excluded from membership of the local medical society (Starr 1982). Doctors working for companies earned the contempt of their colleagues. Insofar as employees wanted free choice of doctor, such choice might have been provided, but other local doctors were unwilling to take patients on the terms offered.

Parallel with company developments were the friendly societies (lodges, fraternal orders, and mutual benefit societies) which became established—often with the same names as their European counterparts. These consumer-run organizations in the 1890s and first decade of the twentieth century contracted doctors for out-of-hospital services on a capitation basis. The societies paid doctors $1 to $2 per member per year, which were regarded by the profession as very low rates (Starr 1982). But estimates of the average earnings of doctors at that time varied from $730 to $1,500 per year (Starr 1982). Thus, at only one dollar per member, a doctor could do well with the care of 2,500 members and very well if the capitation rate was $2 per head. As in England the process of underbidding between doctors appeared, which led to criticism from the *Journal of the American Medical Association* of the "ruthless competition" it "invariably" introduced.

Nevertheless, this type of practice did develop on a considerable scale, particularly among immigrant communities. In New York City there were thousands of funds—mostly branches of larger fraternal organizations. It was estimated that one-third of the Jews in Providence had contract doctors, and in some industrial areas the proportion was as high as 50 percent (Mathews 1909). In Buffalo, New York, lodge practice covered 150,000 persons. The system was also to be found in Pennsylvania, Michigan, Illinois, and California (Starr 1982).

Starr argues that the sharp decline in this type of practice in the second and third decades of the twentieth century was due to "the declining supply of physicians" which "reduced the availability of cheap labor. . . . Doctors could not be found to work on the old terms." This explanation is unconvincing. He quotes that the average

net earnings of doctors in California were $6,700 in 1929 and fell to $3,600 in 1933. The average earnings of workers also rose in this twenty-year period. If $2 a year could be afforded in 1910, $3 could readily be afforded in 1929, and this again would have given a satisfactory living to a young doctor with 2,500 members plus some ordinary private practice.

It seems much more likely that the American Medical Association (AMA) was able to stamp out nearly all of this type of practice by sanctioning doctors who engaged in it. The opposition to this type of practice expressed in the *Journal of the American Medical Association* in 1907 allowed no exceptions. There is "no economic excuse or justification for this sort of practice." The early HMOs existed "for the purpose of buying medical services at wholesale rates and selling them to its members at retail rates. . . . If one doctor agrees to do lodge practice for $2 per member per year, some cheaper men will soon agree to do the same work for $1.50. In this way the price of medicine is reduced to a ridiculously small amount." This stance may be compared with that of the committee appointed by the British Medical Association two years earlier. For all the detailed criticisms of how contract practice operated in some cases, the committee came to a "conclusion which can hardly be set aside" that "in certain districts under existing conditions certain classes of the community require some provision whereby they can insure by small premiums against the cost of medical attendance" (*British Medical Journal* 1905).

Precisely how the AMA stamped out nearly all of this type of contract practice is not wholly clear. Clearly, membership in the county medical society and the risk of boycott from referral and later from hospital privileges were much more effective sanctions in the United States than in Europe where doctors who worked in hospitals were increasingly separated from doctors who worked in the community. Moreover, the AMA had built itself into a much larger and more powerful organization by 1910 than was the case, for example, in England. Possibly the close association of county medical societies with licensing bodies and their ability to influence legislatures to impose sanctions and requirements on would-be HMOs, as well as ostracizing doctors participating in HMOs, succeeded in crushing them. This cleared the road for the emergence of provider-sponsored health insurance on a retrospective reimbursement basis led by Blue

Cross and Blue Shield. While consumers originally chose the pattern of insurance, it was the American providers who decided what would come to be regarded as "conventional" insurance in the United States.

Conclusions

This brief sketch of international experience leads to the following conclusions. HMOs, both those employing doctors appointed on a salaried basis and contracting selected doctors, did operate under health insurance in Europe. President Nixon reinvented a very old wheel. In all the countries examined, the medical associations in one way or another eliminated virtually all price competition. They fought for and won the same three issues. The first was the right of the patient to have free choice of doctor. This could alternatively be interpreted as the right of all doctors to compete for a share of the health insurance cake. But the crucial point was that it meant that individual insurers lost the right to pick and choose their doctors according to their diligent care of patients, their willingness to receive low pay, and/or their economical use health of resources. Reinforcing the right of free entry for doctors was the establishment of what amounted to tenure—the right to remain in insurance-covered practice unless some very grave breach of defined regulations had been proven before some tribunal.

Backing up the rights of free choice and free entry was the establishment of central, regional, or local negotiation of standard terms of payment for doctors. This prevented each insurer making separate deals with selected providers. In other words, it removed price competition. It also made it seem wasteful to have more than one local insurer each incurring administrative costs. Competition had become functionless.

The third issue was the preservation of price discrimination according to the income group of the patient. The medical profession has argued over the years that health insurers obtained medical services at below market prices and these special low prices should not be available to the higher-income groups. Health insurers, on the other hand, have argued that the profession judges market prices by what it attempts to charge private patients. The introduction of health insurance by greatly widening the market and protecting the profession

from bad debts has augmented the income of doctors far beyond what could be obtained in a free market. While the profession prices its workload, health insurers look at the impact of its payments on medical incomes. Naturally these two perspectives cannot be reconciled. But, in practice, price discrimination has been retained in one or more of four ways:

1. The establishment of an income maximum for membership of statutory regulated health insurance, as in Germany, The Netherlands, or the United Kingdom (1911 to 1948).
2. The establishment of wholly separate health insurance funds for the higher-income groups, which yield doctors a higher income for the same services, as in The Netherlands, Germany, and Denmark (1915 to 1976).
3. Reimbursement at standard rates, allowing the doctor or designated doctors to charge more, as in France and Australia (from 1953), and also in New Zealand, Sweden, and, until recently, in some provinces in Canada.
4. Extra voluntary private insurance, which enables patients willing to pay to go to doctors who charge higher than the standard fees, as in Switzerland.

The first two conditions prevent each separate health insurer from finding the best buy in a competitive market. The third condition prevents the whole population from having insurance on the same terms. Free choice is further restricted where there are separate funds for different industries or occupations. Naturally low-health-risk industries or occupations want to preserve their lower contribution for the same benefits or their wider range of benefits paid for by the same contribution.

What lessons do all this have for the future of HMOs in the United States? HMOs now cover only about 12 percent of the population. It was not until European HMOs were covering some 20 percent of the population that the medical profession took up its militant stance about free entry, negotiated terms of service covering all sick funds, and the exclusion of the higher-income groups. If HMOs continue to increase their market share and if an excess of underemployed doctors in the United States accumulates over the coming years, some HMOs may come to operate on the more ruthless private-market

principles of nineteenth-century European HMOs, such as putting jobs out to tender (or, at least advertising them at very low rates) and more readily dismissing doctors who do not keep within the tight limits laid down for the cost of benefits. Possibly a separate HMO doctors' union could emerge. Might there at some stage be a strike or threatened national strike of HMO doctors, calling for standard terms of service, income limits for HMO patients, the right to supplement their incomes from non-HMO patients, and tribunals or courts to adjudicate issues of dismissal? And if there was, could state legislatures and Congress stand idly on the sidelines while 12 to 20 percent of the American population were threatened with either being denied health care or having to pay a further premium to another insurer?

But history seldom repeats itself in precisely the same form. HMOs have not shown themselves anxious to become aggressively competitive. Even if they were, and coverage grew to the extent that some of their more enthusiastic advocates hope, it is by no means certain that HMOs would survive. At least in Europe and Australia the Enthoven competitive market was destroyed once it grew large enough to be a matter of major professional concern. Academics in Europe who advocate competitive HMOs as the solution to both efficiency and cost containment in the medical market need to ponder carefully whether this solution which proved not to be politically viable earlier would survive the test of time if re-introduced today.

References

Abel-Smith, B. 1964. *The Hospitals 1800-1948.* London: Heinemann.

———. 1985. *Eurocare.* Basel: Health Economics.

Akin, J., N. Birdsall, and D. de Ferranti. 1987. *Financing Health Services in Developing Countries.* Washington: World Bank.

Aubert, A.B. 1970. The Norwegian Medical Association As It Should Be. *Tidsskrift for Den Norske Laegeforening* 10b (30 May):51.

Berg, O. 1986. Verdier og Interesser: Den Norske Laegeforenings Fremvekot og Utvikling. In *Legeme of Samfurnet,* ed. L. Oeirind, O. Berg, and F. Hodne, 226. Oslo: Den Norske Laegeforening.

Bjercke, O. 1970. The Norwegian Medical Association As It Is. *Tidsskrift for Den Norske Laegeforening* 10b (30 May):48.

British Medical Journal. 1905. An investigation into the Economic Conditions of Contract Medical Practice in the United Kingdom. 22 July:3, 20–22, 25.

Brown, L.D. 1983. *Politics and Health Organization: HMOs as Federal Policy.* Washington: Brookings Institute.

Green, D.G. 1985. *Working Class Patients and the Medical Establishment.* Oxford: Pergamon.

Hanson, H.C. 1974. *Historien on Sygekasserne.* Aalborg: Aksel Schoelins.

Harris, R.W. 1946. *National Health Insurance 1911–46.* London: Allen & Unwin.

Hogarth, J. 1963. *The Payment of the General Practitioner.* Oxford: Pergamon.

Journal of the American Medical Association. 1907 (Editorial). 49:2028–29.

———. 1911 (Editorial). 57:145–46.

Juffermans, P. 1982. *Staat en Gezondheidszorg in Nederland.* Nijmegen: Sun.

League of Nations. 1924. *Health Organisation in Denmark.* Geneva.

Levy, H. 1944. *National Health Insurance.* Cambridge: Cambridge University Press.

Mathews, G.S. 1909. Contract Practice in Rhode Island. 599. *Bulletin of the American Academy of Medicine* 10:599.

Maurer, A. 1983. Switzerland. In *The Evolution of Social Insurance. 1881–1981,* ed. P.A. Kohler and H.F. Zacher, 384–442. London: Frances Pinter.

Rasmussen, O. 1979. *Sygesikringsloven.* Copenhagen: Juristforbundets Forlag.

Royal Commission on Health Services. 1964. 389. Ottawa: Queen's Printer.

Royal Commission on the Poor Laws and Relief of Distress 1905–9. 1909. 848. Cd. 4499. London: HMSO.

Saint-Jours, Y. 1983. France. In *The Evolution of Social Insurance 1881–1981,* ed. P.A. Kohler and H.F. Zacher, 105, 119–20. London: Frances Pinter.

Sax, S. 1984. *A Strife of Interests.* Hemel Hempstead: Allen & Unwin.

Starr, P. 1982. *The Social Transformation of American Medicine.* New York: Basic Books.

Urdal, N. 1961. *Sykelrygden i Femti Ar. Historien om en Grunnleggende Sosial Reform.* Oslo: Norges Trygdelasslag.

Van der Hoeven, H.C. 1983. *Om de Macht bij het Fonds.* The Hague: Azido.

Zollner, D. 1983. Germany. In *The Evolution of Social Insurance 1881–1981*, ed. P.A. Kohler and H.F. Zacher, 32, 37, 85. London: Frances Pinter.

Acknowledgments: The author wishes to thank Dr. Alexander Preker for supplying evidence and references for Denmark and Norway, and Erik Heydelberg for evidence and references for the Netherlands.

Address correspondence to: Prof. Brian Abel-Smith, The London School of Economics and Political Science (University of London), A243, Houghton Street, Aldwych, London, WC2A 2AE, England.