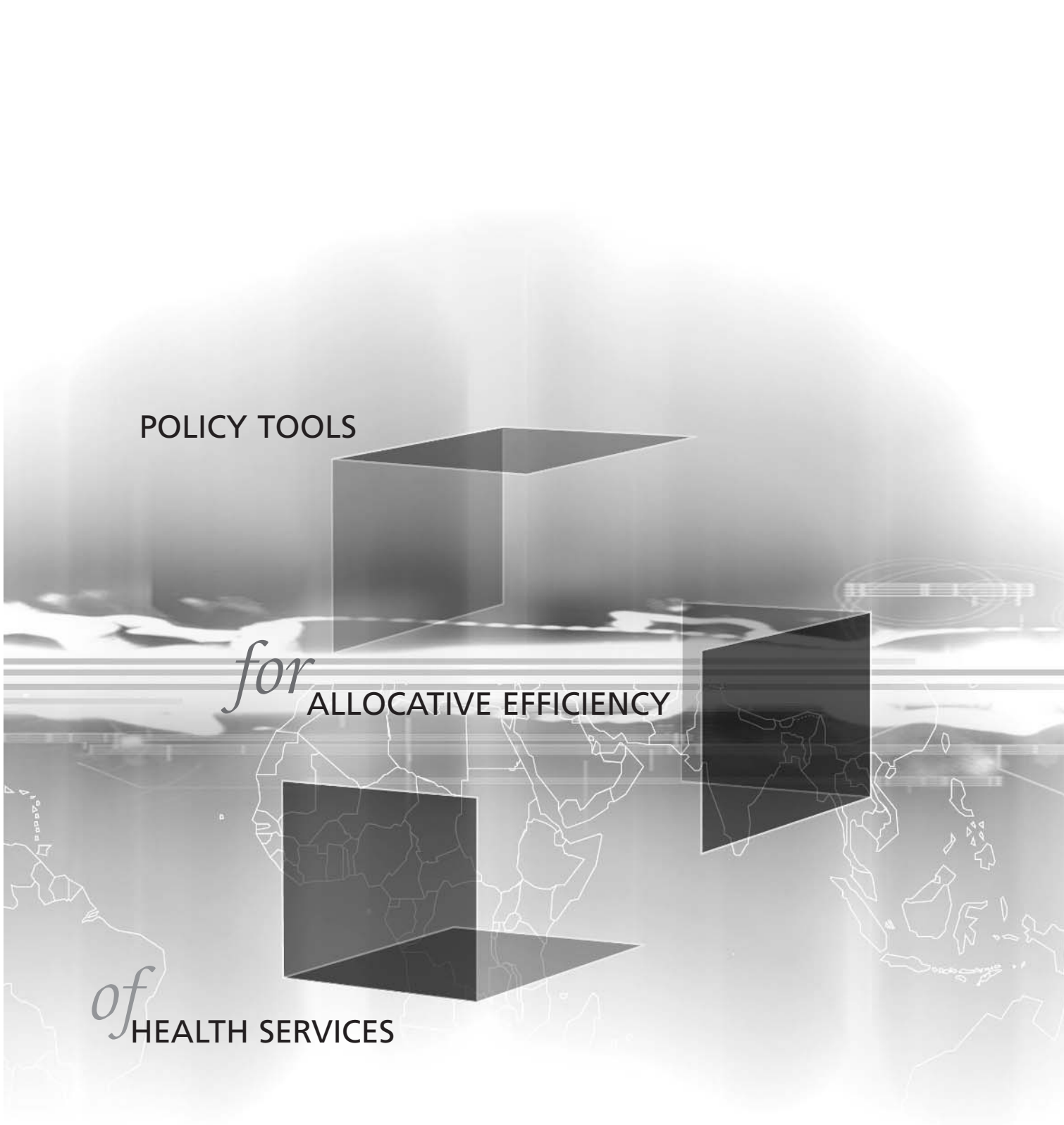
POLICY TOOLS *for* ALLOCATIVE EFFICIENCY *of* HEALTH SERVICES

Xingzhu Liu

WORLD HEALTH ORGANIZATION

POLICY TOOLS *for* ALLOCATIVE EFFICIENCY *of* HEALTH SERVICES



Xingzhu Liu

WORLD HEALTH ORGANIZATION

WHO Library Cataloguing-in-Publication Data

Liu, Xingzhu.
Policy tools for allocative efficiency of health services / Xingzhu Liu

1.Health economics. 2.Health planning - methods. 3.Financial management -methods 4.Policy making 5.Cost-benefit analysis - methods 6.Evidence-based medicine 6.Health care rationing - methods I.Title.

ISBN 92 4 156252 8 (NLM classification: WA 525)

Cover design and layout by James Elrington.

Printed in Switzerland.

Contents

Tables

Figures

Boxes

Foreword

This study by Xingzhu Liu is the result of trying to answer for the World Health Organization the question, "Are there any policy tools that have been shown to improve the allocative efficiency of health care provision?" Two terms in that question, policy tools and allocative efficiency, require some explanation.

A policy tool, as the term is used here, means more than just a policy: it includes some means of implementing or enforcing the policy. Legislation, regulation, accreditation and various kinds of standard-setting, and methods of paying health care providers and of sharing costs with consumers, all fall under this heading. Any policy tool either requires someone to do or not to do something, or provides incentives to do or not to do it. In part, this study is a taxonomy of such tools, which can be classified in various ways: according to who wields them — for example, governments, insurers or professional associations; whom they are used to influence — providers or consumers of health care, or both; what incentives they create; or what effects they can be expected to have. All the policy tools analysed are meant to affect which health care services are actually provided, by whom and for whom.

As economists use the term, allocative efficiency is sometimes translated as "doing the right things"; that is, producing the correct or ideal mixture of goods and services. It is distinguished from technical efficiency, or "doing things right", which means that whatever is produced should be produced with the least possible use of resources or inputs. No more of any input should be used than is necessary; moreover, when several different inputs are combined, and can be substituted for one another at least to some extent, technical efficiency means choosing that combination which produces the desired output at the lowest total cost. What that cost is, depends on the prices of the inputs: when doctors are paid only a little more than nurses, technical efficiency may mean using relatively many physicians and few nurses, whereas when doctors' pay is much higher, it becomes efficient to have nurses do as much as possible of the work that physicians might do. In both cases, neither more doctors nor more nurses should be employed than is necessary.

Technical efficiency by itself is of little value, since it is possible to produce "the wrong things" and still waste no resources in the production of each one considered separately. Allocative efficiency means that, in addition, the output that is actually produced could not be changed — more of some good or service and less of another — without making some consumer of the goods and services worse off. Gains for some could come only at the cost of losses to others, so there would be no room for changes that everyone could agree to. Under the conditions of a perfectly competitive market, efficiency in "the right things" to be produced can be identified with "those things consumers want, and are willing to pay for". That is, allocative efficiency means satisfying consumer demand, given the distribution of tastes or relative valuation of different outputs and the distribution of income (which also depends on who produces what). It is presumed that what consumers demand will reflect, among other things, any needs that they may have, so the latter are not taken explicitly into account. The equity of the resulting distribution of goods and services among different people is similarly not considered.

These suppositions are widely regarded as particularly inappropriate in the market for health care services, which is characterized by more "market failure" than are most other sectors of the economy. So far as efficiency is concerned, people often have needs which they do not recognize and so do not translate into demand; they may also demand, or be led by providers

to demand, services for which there is little or no need. Besides the difficulty of judging needs, consumers have an especially hard time judging the quality of health care. Many wants and needs are not translated into demand because people cannot afford to buy the corresponding services out of pocket; or cannot afford, or do not have access to, adequate insurance against the financial risks posed by accident or illness. And private, competitive voluntary insurance is also a source of market failure, even apart from differences in income, because of differences in health risks and in people's knowledge about them. For these reasons, "the right things" to produce in health care cannot simply be identified with those things that people want and can pay for by themselves. Allocative efficiency, it appears, must be given some other meaning.

One alternative meaning, which is used in this study, relates the costs of different health care services to the health outcomes — the improvements in health — they yield. It does not matter, for this purpose, whether the improved health results from preventing a loss of health from disease or accident, restoring health or curing the health problem, or palliating its consequences. First one must define a metric for health status, such as years of life saved, which may be adjusted for quality of life or for disability. Then any health care service or intervention is in principle characterized by both a health outcome and a cost of provision, both of which may vary with the scale of production and with the characteristics of the providers and the consumers. The ratio of cost to outcome is then the cost-effectiveness ratio of the service or intervention, estimated as an overall average or for larger or smaller changes at the margin. Allocative efficiency results, on this definition, from that combination of health care interventions which, besides minimizing the cost of producing each service, maximizes cost-effectiveness. That is the combination which delivers the largest gain in health status for a given total expenditure; or requires the smallest expenditure for a given improvement in health. This is what is usually meant by "getting value for money" in health care.

This notion of allocative efficiency requires three assumptions. First, it is only the health outcome of care that matters: comfort or amenities while undergoing treatment, relief of doubt or anxiety, the personal relations with one's care provider, the degree of financial risk or of protection from it, and any other consequences of consuming health care are not considered part of the output. (The consequences or benefits for the providers are also not taken into account.) Second, a unit of health improvement is equally valuable, no matter who benefits from it. Depending on how health status is measured, there may of course be differences among individuals; for example, an extra year of life may be thought to be worth more at younger than at older ages. However, once the metric is determined, the object is to maximize the total health gain without considering how that is distributed among the users of health care. Third, the resources needed to provide care can be transferred from one user to another, in pursuit of that maximum health gain, without any cost other than the cost of the care. That is, if one patient has received care to the point where an additional dollar's worth of resources would yield less gain than if it were spent on a different patient, then the resources can be shifted to that second patient without any frictional or transaction costs.

These assumptions come closest to being satisfied when essentially all health care is controlled by a single agency such as a national government, which pays for the care even if it is not the direct provider, and which is equally responsible, and responsive, to all citizens. Under those conditions, maximizing cost-effectiveness may be a reasonable policy objective, so it is also reasonable to search for tools which may promote that outcome. As there are more and more-agencies paying for health care and making decisions about its distribution, these assumptions become less realistic. There are two strong reasons for this. First, an individual paying for his or her own health care will not necessarily want or demand the most cost-effective intervention for his or her particular health problem. The additional improvement in health to be had from

a somewhat less cost-effective service may seem well worth the extra cost, given the person's capacity to pay and his or her appreciation of the value of the other uses for the same money. Both costs and effects matter, but decisions may not be based on the ratio of the two. Second, to choose the most cost-effective interventions for the whole society means, necessarily, to make choices also among individuals, because some people's health problems admit of more cost-effective remedies than do those faced by other people. Thus allocative efficiency in the sense used here means that the last or marginal dollar of resources has the same chance of contributing to improved health, no matter whom it benefits. It does not mean that every individual has the same chance of having his or her health problem resolved.

Given these limitations, which arise from how health systems are organized and not from deficiencies in the analysis, it is not surprising that there is rather little evidence that particular policy tools actually improve allocative efficiency. When the agency employing a tool is not a government but an insurer or other private organization, maximizing health gain may not even be an objective. Controlling total costs, in contrast, is nearly always an objective. Much of the evidence that Dr Liu has assembled concerns ways to control either the costs of individual interventions, or the total expenditure on services for a defined population. This is of course consistent with technical efficiency, which is always desirable. It is also consistent with allocative efficiency, to the extent that it means promoting more cost-effective means of dealing with any particular health problem, whether or not explicit choices are made about the relative priority of different health problems.

Rather than look, in his analysis, at only those policy tools with a clear impact on overall cost-effectiveness, Dr Liu has chosen to examine all the available tools for which there is some evidence of an effect on which services are provided, how, by whom and to whom. Many of these tools, even if designed or used primarily for other purposes, may also affect allocative efficiency, and such secondary effects are potentially important. Moreover, tools may be employed in combination, to reinforce desirable effects or to offset or compensate for undesirable effects. From this it follows that health policy-makers need to think not only of which individual tools to develop and use, but of how to combine them in a "toolkit" for best results. This study does not simply tell policy-makers which tools to take up, but rather how to determine which ones are most likely to be feasible and effective, given particular circumstances of political and administrative capacity, knowledge on the part of both providers and consumers, and other factors which affect whether a policy tool can be used, and with what outcome.

This study draws on a very large number of published sources, both theoretical and empirical. It is not surprising that most of these refer to high-income countries, and a very large share of them deal only with experience in the United States. That country has an extremely complex health sector, characterized by a number of both natural and deliberate experiments in the regulation, financing and delivery of health care. The United States not only has the highest health care expenditures in the world, but — partly because of that fact — also dominates health care research to an even greater degree. The variety of research on a number of problems and many different policy tools provides much potentially valuable information. What is less clear is how far the evidence from the United States and a few other high-income countries, notably the United Kingdom, can be used to improve health policy and outcomes in middle- and low-income countries struggling with some of the same problems but with differently organized health systems and with far less capacity, financial or otherwise, to deal with those problems. Dr Liu therefore concludes his study not only with advice on how to judge the value of different tools and put them into a serviceable toolkit in such countries, but also with suggestions for the kinds of research that would be most valuable for pursuing allocative

efficiency in much of the rest of the world. Theory is essential for initial guidance, and actual experience in high-income countries may be of great value; what is also needed is empirical verification, and adaptation as necessary, of the usefulness of different tools in countries where such knowledge is still quite scarce.

Philip Musgrove
Fogarty International Centre
National Institutes of Health

Acknowledgements

This publication is an outcome of the research project sponsored by the Global Programme on Evidence for Health Policy (GPE) at the World Health Organization, and implemented when I worked at WHO, initially as a Global Leadership Fellow in Health and then as a WHO staff member.

I am especially grateful to Dr Christopher Murray, Executive Director of the Cluster of Evidence and Information for Policy (EIP). He not only selected the topic addressed here, but also provided guidance throughout the study. Special thanks go to Ms Kei Kawabata, Coordinator of the FAR group (Designing Policy Options: Financing, Resource Allocation, Regulation and Legislation) within the EIP Cluster, for her leadership, management and technical guidance. The final production process was supported by Richard Poe.

I want to express my deep gratitude to those who provided comments that resulted in both structural changes and technical improvements of the manuscript. They are Dr Guy Carrin, Mr Andrew Creese, Dr David Evans, Dr Neelam Sekhri Feachem, Mr Joseph Kutzin, Dr Yunkap Kwankam, Dr John Langenbrunner, Dr Philip Musgrove, and Dr Tessa Tan-Torres Edejer.

I greatly appreciate Angela Haden's structural editing of the initial and final drafts according to the comments provided by these reviewers, which resulted in a shorter and more tightly organized text. In a subsequent revision, Dr Philip Musgrove reviewed the manuscript carefully and suggested both technical and linguistic improvements. His generous help and substantive editing has greatly improved the quality and conciseness of the manuscript.

Chapter 1. Health Service and Policy Efficiency

Efficiency is a central theme of any sector of the economy, although the means for achieving it differs depending on the characteristics of various sectors. In the health sector, with the same level of health expenditure per capita, the level of health achieved significantly differs among countries (1). Why do some countries perform better than others? Besides other factors that may affect health, such as education, geographical location, genetic and lifestyle differences, the efficient use of available resources in the production of health matters. Based on a review of the literature, this study considers how governments can intervene to achieve health system efficiency, and suggests possible policy tools.

Policy and the need for evidence

The Oxford Dictionary defines policy as "a course or principle of action adopted or proposed by a government, party, business, or individual" (2). Usually, a government policy has the following four characteristics: (1) it corresponds to specific issues which are on the government agenda; (2) it is based on formally articulated decisions that are stated in a legal document; (3) it affects a significant number of people (hence the definition here is equivalent to public policy); (4) it needs rules for operation and guidelines for implementation. The distinction between a policy and a policy tool is that the latter refers to a specified course of action, while the former can be general principles stated formally by authorities.

In health care policy-making in many countries there is rhetorical emphasis on evidence-based policy. The major argument is that a policy as an intervention to deal with some specific problems is less likely to be effective and efficient if the policy is not based on scientific evidence. Muir Gray outlined how evidence from research into financing and organization of health services may be used in developing health care policy (3). However, policies, including health policies, are rarely based on scientific evidence. As recognized by Murray & Lopez (4), while many health policies directly deal with health problems, the international public health community did not routinely quantify or project the health problems of populations; there were no standardized compilations of comparable information on the extent of morbidity, disability and death in different populations of the world; too often advocates of a policy provide filtered or biased information, and health problems without vocal advocates are frequently ignored until policy neglect leads to crises in which the public demands action. The lack of evidence-based health policy is due partly to the lack of scientific evidence when policy-making is urgently needed, partly to the fact that the interaction of stakeholders in the policy-making process often makes the scientific evidence less important (as in the case of the Oregon list of health services, see section 5), and partly to a lack of culture of evidence-based policy among policy-makers. As Ham et al. argued, the creation of a culture of evidence-based policy-making is important for its widespread acceptance and actual implementation (5).

The lack of evidence-based policy does not mean evidence is not needed in policy-making. Based on the work of Walt (6), the policy-making process can be divided into six stages:

- *Stage 1:* Identification of problems — the major activities in stage 1 are to quantify the current problem and project its future development; this stage is an evidence generation process through which the size of the problem is measured.

- Stage 2: Establishment of a policy agenda — the size of the problem that is specified and quantified in the first stage of the policy-making process is a key determinant of whether the problem will be put onto the policy agenda. In this stage the evidence generated in stage 1 will be used by advocates to mobilize the attention of policy-makers and by the latter to make decisions.
- Stage 3: Specification of policy alternatives — in formulating the alternatives, policy-makers and researchers may use international experience and evidence, summarize the experience of their own countries, conduct experiments to generate evidence for the alternatives, or simply formulate alternatives based on theories which have been generalized from the synthesized evidence. In this stage, evidence on the feasibility, effectiveness and cost of each of the policy alternatives must be generated or provided.
- Stage 4: Choice among the specified policy alternatives is actually an evidence utilization process, in which the key policy-makers and stakeholders compare the evidence for the available alternatives and make final choices.
- Stage 5: Implementation includes the detailed design and practice of the policy; it serves as a natural quasi-experiment, through which further evidence on the effectiveness, costs and implementation problems will be generated.
- Stage 6: Evaluation of the effectiveness of the policy — evidence generated in the implementation of the policy can be collected, analysed, and compared with what is projected. This is also an evidence utilization process in which information about success and failure can be used to feed back to the design and implementation of the policy.

Generation and utilization of evidence is related to every stage of policy-making process. The evidence can be generally divided into three types: on the need for policies (related to stages 1 and 2), on policy options (related to stages 3 and 4), and on policy effectiveness (related to stages 5 and 6). The focus here is on the last type of evidence, concerning the effectiveness of policies that have been implemented in various countries.

Ideal resources allocation

When talking about improving allocative efficiency, one approach is to set up an ideal scenario in which resources are allocated in an optimal way, and then to examine the possible policy tools to move towards this ideal situation. To proceed, we first examine the process of health production in which resources are allocated; then set up ideal scenarios of resource allocation; and then specify the policy "anchoring points" where policy tools can be located to promote allocative efficiency.

Health production is divided into three stages: the production of factors; the production of health interventions; and the production of health. The outputs of intervention production are inputs to health production.

In health production, a number of resource allocations affect allocative efficiency:

1) the allocation of funds for purchasing different inputs to the production of health, such as workers, buildings, and equipment;
2) the allocation of these inputs among various factor outputs by the planners of factor production;

3) the allocation of funds for purchasing various factors to use as inputs to produce health interventions by a health care provider;
4) the allocation of intervention inputs among various types of health interventions (including geographical areas, types of health facilities);
5) the allocation of intervention inputs by a health care user to produce health;
6) the allocation of health interventions among users with different health problems and conditions for the production of population health by health intervention providers and planners.

To achieve allocative efficiency of the whole process of health production requires that resources are allocated optimally for each of these six steps.

Scenario 1: In the production of a factor, the producer allocates its total funds among various inputs in such a way that the inputs are combined for the production of a given amount of the factor at least cost. For example, if there are only two factor inputs for the production of medical doctors — labour and capital — efficiency requires them to be combined in such a way that the cost for producing a given amount of doctors is minimized. Too much capital and too little labour means a departure from this ideal scenario. However, efficiency of factor production does not guarantee the best combination of factor outputs, which is another requirement of efficiency of health production.

Scenario 2: The funds are allocated among different factors (doctors, nurses, medical equipment, medical materials, etc.) in such a way that the best mix of factors can be produced, not only reflecting the need of society for health interventions, but also that the cost for producing health interventions is minimized. If there are only two factors needed for the production of health services — doctors and nurses — efficiency requires that the numbers of doctors and nurses are combined in such a way that the cost for producing a given amount of health services is minimized. Too much funding for the production of doctors and too little funds for the production of nurses reflects a departure from this ideal scenario.

Scenario 3: To produce health interventions, a health care provider allocates its funds among various inputs to produce a given amount of health intervention at least cost. However, allocative efficiency of intervention inputs by health care providers does not guarantee that health care providers produce the best mix of health interventions; that is a necessary but not sufficient condition for the production of the best mix of health interventions.

Scenario 4: The production of the best mix of health interventions requires that intervention inputs be allocated appropriately amongst the types of interventions. The stock of intervention inputs must be allocated among geographical areas (e.g., urban versus rural) and types of health facilities (e.g., curative versus preventive; tertiary versus primary), because the location of intervention inputs, to a large extent, determines the types of intervention outputs. For example, imbalance of doctors between urban and rural areas may result in overprovision of cost-ineffective interventions in the urban area and underprovision of cost-effective interventions in the rural area; too few human resources in disease prevention may lead to underprovision of preventive care.

Scenario 5: In the production of individual health, to achieve allocative efficiency, the purchasers (users or representatives of users) use their funds to purchase the best mix of health

inputs so that the cost of health inputs is minimized for achieving a given level of health gain. The users of health interventions choose from among various health inputs for the improvement of health depending on their health conditions: healthy foods, physical exercise, immunization, doctor visits, hospitalization, surgery, etc. Because the users lack information for making rational choices of health inputs, much of the decision is usually delegated to health care providers. In addition, third party (insurers) are increasingly involved in purchasing health interventions on behalf of patients, particularly in managed care where insurers often make decisions on what types of services should be provided and under what conditions.

Scenario 6: The health interventions produced by health care providers are delivered to the people who need them the most, and thus yield the maximum amount of health gain. Producing the best mix of health interventions is meaningless, unless these interventions are delivered to the right people at the right time. For example, the production of vaccines and immunization services will improve the mix of health interventions; but unless they are delivered to the target population (e.g., children) before the epidemic period of the year, the impact may not be great. Improving the intervention mix may not achieve the best result in health improvement.

The overall allocative efficiency of the health sector requires the simultaneous satisfaction of the six scenarios in which health resources are perfectly allocated for each step and each type of production. To sum up, the allocative efficiency of health resources requires:

- the right inputs for the production of a factor (scenario 1);
- the right factor outputs (scenario 2);
- the right inputs for the production of an intervention (scenario 3);
- the right intervention outputs (scenario 4);
- the right health inputs for the production of an individual's health (scenario 5);
- the right health inputs to the right people (scenario 6).

The six scenarios of resource allocation are not independent of each other. For example, the right factor outputs are constrained by having the right inputs for the production of factors; the right intervention outputs depend on having the right inputs for the production of health interventions; and distorted intervention inputs will result in distorted intervention outputs. The implication of these scenarios and their relationships is that to improve the allocative efficiency of the health sector, policy tools are needed corresponding to each of the six types of resource allocation.

Allocative efficiency of health interventions

Health system performance is concerned not only with the degree of achievement of the primary goals of the health system (such as health status of the population, responsiveness to their legitimate expectations and fairness in financing health care), but also the cost at which the goals are achieved. Efficient performance requires that the level of goal achievement is maximized for a given level of cost, or the cost is minimized for a given level of achievement.

While efficiency of health system performance requires both technical and allocative efficiency for the achievement of all goals, this paper focuses on allocative efficiency for the achievement

of health. The concept of allocative efficiency takes health interventions (including services, drugs, and other activities, the primary intention of which is to improve health) as inputs and heath of the population as an output. In other words, allocative efficiency refers to the maximization of health outcome with the least costly mix of health interventions.

Theoretically, such a mix of health interventions must be provided so that the marginal health outcomes of a unit of monetary input of various interventions are equal. Practically, to improve allocative efficiency, resources should always be used to produce the most cost-effective interventions. The cost-effectiveness of health interventions has three dimensions. First, the providers of health interventions have to produce a mix of health interventions that reflects people's need and has a potential to yield the highest return on health. Second, each unit of the produced health interventions must be delivered to and used by the people who need them the most and can get the maximum health gain. Third, health interventions are provided at the least costly location. As Meekison simply put it, allocative efficiency of health interventions must ensure the right care is provided to the right people at the right place (7).

Evidence of allocative inefficiency

Reality is far from this ideal, however. Achievement of similar levels of health outcomes with a wide variation in total resource use among countries points to the possibility of containing or reducing expenditures by a better allocation of health resources, while still maintaining favourable health outcomes. When the experiences of the OECD countries are compared, it becomes obvious that health care outcomes are not very sensitive to variations in health care expenditures (8). Life expectancy and infant mortality measures are similar in OECD counties, but, in contrast, the variations in total health care expenditures are very large. Studies in eight OECD countries show that although there is little difference in the health status of the population, there are large differences in health care expenditure, resources allocation, and use of services. For example, the ratio of physicians to population ranges from 1.4 to 3.1 per 1000; the average number of physician visits per person per year ranges from 2.8 to 11.5; the number of hospital beds ranges from 4.7 to 12.4 per 1000 population; the annual number of hospital days per person ranges from 1.2 to 3.7; the number of MRI scanners per million population ranges from only 1 in Canada to 11.3 in the United States; and total health care expenditure as a percentage of GDP varies from only 6.5% in Denmark to 14% in the United States of America (9).

Although there is evidence that allocation of more resources to primary and preventive care, which is highly cost-effective or even cost-saving, will improve allocative efficiency, there seems to be no clear mechanism for achieving the desirable resource shifts. In the United States, preventable illnesses make up approximately 70% of the burden of illness and of the associated costs (10). Disease prevention is proved to be a far more comprehensive and compelling solution to improve population health and control costs, and the experiences of many prevention programmes have documented reduction in the costs of expensive medical interventions of 15–20% and even more. Nevertheless, in 1994, the per capita public health expenditure for disease prevention and control in the United States was only 3% of the national health expenditure. In Australia, epidemiological studies suggest that many of the health gains of recent decades have been a result of public health initiatives. Despite this, in 1994–95 only 5% of Australia's health budget was allocated to public and community health and less than half of this for prevention (11). In low- and middle-income countries, although an

essential package of cost-effective health interventions can eliminate about 30% of the DALY lost at a little cost, these health interventions are underfunded and resource allocation is biased towards less cost-effective hospital services (12).

There is evidence that medical services and drugs are often provided inappropriately, with much variation in provision of surgical procedures among small areas in the United States. Wennberg & Gittelsohn found that the number of procedures performed per 10 000 persons ranged from 13 to 151 for tonsillectomy, 10 to 32 for appendectomy, and 30 to 1141 for dilation and curettage (13). Studies also show that physicians vary greatly in the rates at which they order laboratory tests and X-rays: the highest users may order tests 10 to 20 times as often as the lowest users. The rate of test use does not appear to have a consistent relation with the outcome of care (14). In the United States, studies have shown that 14% of the coronary artery bypass surgery (15), 20% of the heart pacemakers (16), 20% of hospital days (17) and 33% of carotid endarterectomies (18) are unnecessary or inappropriate. In Germany, drug expenditure accounted for 19% of total health expenditure (19). Surveys indicated that doctors in western Germany prescribed on average about eleven medicines per person per year, almost three times more than American physicians. German doctors admitted that 40% of the prescriptions were unnecessary, ineffective, or even harmful. In Canada, it was estimated that 15–30% of a wide range of services were provided inappropriately (20), and four studies of adult patients showed that from 24% to 90% of admissions and from 27% to 66% of days of hospital stay were inappropriate because of absolutely unnecessary care (2% of admissions and 7% of days of stay) or failure to provide alternative care with a lower level of cost for the same level of health outcome (21). In low- and middle-income countries, it was estimated that improving drug prescribing could save up to 50–70% of national expenditure for drugs without affecting health outcome (22).

This evidence shows that resources are often misallocated — going to less cost-effective health interventions when more cost-effective interventions are available; or to those interventions which yield no improvement in health outcome. Obviously, by improving allocative efficiency of health interventions, the level of health outcome can be improved with the same amount of health care resources; or for the given level of health, the use of health resource can be reduced, for both low-income and high-income countries.

Policy tools for improving allocative efficiency

Allocative efficiency can hardly be achieved through traditional market mechanisms because of the well-recognized market failures in the health sector (23). While there are arguments that market mechanism should not be abandoned and can be used in conjunction with government intervention to achieve economic efficiency (24-26), this study does not focus on the effect of market mechanisms. Instead, the interest is to review the evidence of the effect of various non-market policy interventions. In the history of modern health systems, the health sector has never been treated just like other sectors which produce private goods and services and where free market approaches are used to achieve economic efficiency. Government interventions are in place in both low-income and high-income countries. However, almost all the tools aim at cost containment; the possible effectiveness of these tools on allocative efficiency is not well understood by policy-makers and researchers; and systematic reviews of scientific evidence on the effect of these tools on allocative efficiency are hardly available.

The objectives of this study are to review the literature on available policy tools with a potential to improve allocative efficiency of health interventions; to analyse the evidence on the successes and failures of these tools; and to provide suggestions for future research related to the improvement of the existing tools and the development of a toolkit.

Organization of this study

The following sections review the possible tools related to efficiency of heath care systems. For each tool, we first provide a definition and some background, followed by information concerning the implementation of the tool in various settings and countries. Where it exists, the evidence of the effectiveness of the policy is reviewed. In reviewing the evidence, we first look at whether the policy can contain health care cost; and then at whether it can improve quality and health outcomes. If a policy tool can neither contain cost nor improve quality of care and health outcomes, the review will stop there because these two pieces of evidence are enough to suggest ineffectiveness on the improvement of allocative efficiency. If a policy can either contain costs or improve quality and health outcomes, or both, we proceed to examine the magnitude of efficiency improvement and to look at whether the improvement is a result of changing the mix of interventions, namely, providing more cost-effective health interventions rather than less cost-ineffective ones. For those policies which have negative or no effects on costs, quality or efficiency, we explore the reasons.

The policy tools reviewed are as follows.

- *Section 2:* health resources planning, which includes budgeting inputs for public provision, funding inputs for private provision, controlling capital investment, and technology regulation (including technology assessment and technology control);
- *Section 3:* economic incentives for providers, which includes ways of paying health care providers (fee-for-service, capitation, salary, daily payment, line budget, global budget, and performance-related pay), rate setting, fee structuring, separation of drug dispensing from prescribing, and payer integration;
- *Section 4:* economic incentives for users, which includes user fees and cost sharing;
- *Section 5:* defining benefit packages, which includes rationing, and specifying essential packages of services and essential drug lists;
- *Section 6:* informing providers, which includes medical practice guidelines, use of a prescription formulary, and utilization review.

Section 7 deals with managed care, which is de facto a combination of various tools. Section 8 provides a summary of policy tools. Section 9 makes recommendations for a policy "toolkit", a group of tools that together can improve the allocative efficiency of health interventions. In conclusion, section 10 makes recommendations for policy and future research.

Chapter 2. Health Resources Planning

Resources planning is not a new concept and can be dated back to 1920, when the Dawson report was published in England. In the former USSR, regionalization of health services was introduced progressively as a logical consequence of the regional outlook of planning. In 1933, Chile was divided into twelve hospital zones, for which hospitals of different levels were planned. It was, however, only in the 1940s that, partly as a consequence of the increase of specialized care, regionalization was enforced in many countries. France (Hospital Law of 1941), the United Kingdom (National Health Service Act of 1946) and the United States (Hill-Burton Law of 1947) adopted the principle and were followed by many other countries (27). In 1957, the World Health Organization formally promoted the practice of health resources planning (28). Since then, health resources planning has become an important regulatory activity in many countries.

Policy tools for resources planning vary between countries with a more centrally planned health system (e.g., most European countries and socialist countries before their transition) and those with a more market-oriented health system (e.g., the United States). Most of the planned health systems follow a pattern of public provision, in which the location of health facilities of different levels, the amount of capital and human resources, and the funds for acquiring inputs by facilities are decided by governments. A major tool in this type of system is budgeting inputs for public provision. In the market-oriented health systems, the provision of health care is dominated by private entities, and government funding can hardly influence the overall resources allocation of a country. Various policy tools have been used to lever resources allocation, including funding input for private provision (e.g., the Hill-Burton Program and Health Service Corps in the United States), Certificate-of-Need (CON), capital expenditure cap, moratoria policy in the United States, and technology regulation. Because private provision exists in planned systems and public provision in market-oriented systems, the tools are actually mixed between these two types of systems.

Budgeting input for public provision

How much a government spends on health care will affect the total cost of health care in a country; and where government budgets for provision go affects the location of providers and their services, as well as the types of services provided, and hence the access to care, the quality of care, the health status of the population, and the allocative efficiency of health interventions. In countries where a government funded public provision system dominates, such as Australia, the United Kingdom, and most other European countries, the role of budgeting input for public provision is fundamental. In countries where provision of health interventions is mainly private, governments often play an active role in the provision of community health services and medical education. For example, in the United States, where there is the most privatized health care market, there is also a big public delivery system. This system included 90 urban public hospitals owned by city or county governments, 45 state-owned university hospitals, and 1770 community hospitals in the early 1980s (29). In Japan, the government provides funds to construct elder-care facilities and contracts out for their operation. Although health care is predominantly provided by the private sector in some countries, government has some public facilities to provide interventions (e.g., health services that are public goods) that are not provided by the private sector (30).

Theoretically, under the publicly funded provision system, the allocation of health resources is under the control of government. Government should be able to determine the bulk of spending on health care, and what and where health interventions are provided. Government should be able to direct limited resources to the interventions that are the most cost-effective. For example, the government budget can be directed to priority geographical areas (e.g., the underserved rural), cost-effective service domains (e.g., preventive and primary care), and the organization of programmes that target special diseases (such as the TB control programme in the People's Republic of China, the malaria control programme in Thailand, and immunization programmes in almost all countries). In this case, there is a greater likelihood that government can control the total cost of health care, and determine the access and quality of care; and that the right health interventions are provided to the right people at the right place to improve the health of the population and allocative efficiency.

While it can be observed that — in all countries where there is a dominant public provision system — the total cost of heath care has been kept at an acceptable level, there is no literature analysing whether there is a causal association between government provision and cost control. While public provision has been proved to be effective in improving access to care and the level of health of the population in China and the former Soviet Union, the same system seems not to work in most African countries. The relation of public provision with the access and quality of care, and the level of health is not well documented.

Although government is in a good position to improve the allocative efficiency of health interventions, experiences in many countries show that this is not achieved in practice, either because the right budget policies are not in place, or because the right policies are not effectively implemented. The fact is that government funding is usually concentrated on urban areas, curative services, and hospital services; and is much lower for rural areas, preventive services, and primary care. There is usually overprovision of less cost-effective health interventions and underprovision of more cost-effective health interventions (31). For example, in the high-income countries with public provision systems, the government budget for health promotion and disease prevention is less than 2% of the total health budget, and this percentage shows little increase over the last 20 years (32). In low- and middle-income countries, public hospitals may absorb 60–80% of government health expenditure, and the lion's share of this expenditure is often absorbed by tertiary and secondary hospitals in urban areas (31).

There are two major explanations commonly offered for this situation. First, policy-makers often have difficulties making decisions based on technical analysis — the ministries of health lack information on both cost and effectiveness of health interventions, and the allocation of government budgets is often the outcome of political considerations by different government bodies (31). Second, public choice theory explains much of the misallocation: more influential consumers and producer groups are able to divert resources to the overprovision of costly services that predominantly benefit upper-income groups and have a much lower social rate of return, at the expense of providing basic services to the underserved (33).

For example, although the Chinese Ministry of Health adopted a policy in favour of preventive, primary and rural health care, there has been little practical implementation over the past two decades, leading to deterioration of public institutions providing these cost-effective health interventions, such as township health centres and anti-epidemic stations. The failure of efficient budget allocation of the government funded provision system, plus the well-recognized low productive efficiency of publicly funded health facilities have led to reforms that privatize or provide autonomy to the publicly funded institutions in low- and middle-income countries (e.g., in China and most African countries); and to a provider–purchaser split

in high-income countries (e.g., Australia, the Netherlands, New Zealand and the United Kingdom). The aims of these reforms are to increase the financial accountability and productive efficiency of public facilities, and to channel limited government funds to cost-effective health interventions.

The primary objective of budgeting input for public provision is to increase access to care. Theoretically, the quality of care and health status of the population can be improved through improvement in access to care; and the level of cost, the pattern of utilization and the mix of interventions are well under the control of the government if public provision dominates the health care system. However, the effects of budgeting input for public provision on cost, quality and health outcome are not well documented; and there is substantial evidence that government funds are misallocated.

Funding inputs for private provision

In countries where private provision dominates, governments usually take active roles in directly funding private providers, when there is a shortage of supply or imbalance of providers. The activities or programmes include strengthening medical education, strengthening health facilities, training grant incentives for education of rural physicians, and funding the production of cost-effective products.

Strengthening medical education

Medical education determines the stock and mix of health professionals, which then will affect the access and the types of services provided to the population. Oversupply of medical specialists is likely to lead to overprovision of specialist services; undersupply of primary doctors will limit the access to cost-effective preventive and curative primary care. Shortage of nurses is likely to lead to low quality of nursing services. Theoretically, policy interventions to increase the supply of health professionals in prominent shortage will help to increase access, and improve quality and the mix of health interventions. This is one of the major reasons why governments in various countries, where private provision dominates the medical market, still provide financial support for the education of some types of health professionals.

Although government funding of medical education may occur in every country, evidence of the impact of this policy on access, quality, health and allocative efficiency is rare; and the only available evidence can be found in the United States. In the period of general shortage of physicians at the end of the 1960s in the United States, the government provided substantial grants to construct, expand and maintain medical schools (34). Federal assistance for health manpower training began on a large scale in 1963 with the Health Professions Educational Assistance Act. This Act was amended in 1965 and renewed in 1968 and 1971. The direct recipients of such federal assistance have been both schools (including public and private) and students in the health professions of medicine, dentistry, osteopathy, podiatry, pharmacy, optometry, veterinary medicine and public health (35). In addition, the United States government supported graduate medical education through the Medicare programme by providing funds to teaching hospitals to reimburse them for the cost of teaching.

This increased the supply of doctors, particularly in the inner-city hospitals which rely heavily on the residents graduated from foreign medical schools and which provide services mostly to the poor. These programmes effectively increased the supply of medical doctors and their services in the 1960s and 1970s. Apparently the access to care was improved as a result of the marked increase in the number of physicians. However, whether quality of care was improved is less documented; and whether the health intervention mix changed would depend on what types of physicians were educated.

The United States currently does not have a national policy to regulate the number, speciality choice, or distribution of its physicians. Experience shows that the continuous support for medical education has not only led to oversupply of physicians, but worsened the mix. According to McEldowney et al., in 1992, a total of 62% of United States physicians were listed as specialists and only 38% were in primary care; there was an oversupply of specialists of more than 60% (36). As a result, the health intervention mix is actually worsening (37): too little of more cost-effective primary care and too much of less cost-effective specialist services. Although the distortion of intervention mix can be attributed to many causes, one of the reasons is the imbalance of the types of clinicians as a result of the poorly planned medical education and health manpower programmes.

Strengthening health facilities

In a medical market dominated by private provision of medical services, government does not look on with folded arms. Besides regulation of private provision, governments often provide financial support to private facilities when it is believed necessary for improving access and quality, and correcting an imbalance of providers.

A typical example is the United States Hill-Burton Program. The Hospital Survey and Construction Act, commonly known as the Hill-Burton Program, was designed to assist construction and modernization of American hospitals after the Second World War. The purpose of this programme is to increase supply of and access to care; redistribute hospital beds between urban and rural areas; and redistribute physicians by attracting more doctors through construction of hospitals in underserved areas. The surgeon general was authorized to issue regulations requiring the hospitals to provide uncompensated care and community service as a condition of the receipt of Hill-Burton funds (totalling US$ 6.5 billion).

This programme gave priority to states with lower incomes, rural areas and communities without health facilities. Studies showed that the Hill-Burton Program greatly increased the access to care by residents, particularly the poor; it had a significant effect on the redistribution of beds and physicians between and within states (38), but there were also claims that because of a lack of effective enforcement, there was widespread noncompliance by hospitals in providing uncompensated care (39). In general, the Hill-Burton Program increased the number of hospitals and improved patients access to hospital care.

However, by the end of the 1960s, it was recognized that the nation had built too many hospitals and beds; by the end of the 1970s, fears of bed and physician shortage had given way to concern about the mounting physician and bed surplus (40). By the end of the 1980s, the inappropriate increase in the number of physicians had led to concerns about supplier-induced

demand and the provision of unnecessary care (41). By the end of the 1990s, as it was estimated, 30% of hospitals would have to be closed in the United States (42). Although the programme increased access to care, the oversupply of hospitals and beds may have resulted in the overprovision of services that should have been provided in non-hospital settings.

One of the lessons from the United States experience is that the implementation of a policy should be monitored over time, otherwise things can be overdone, solving one problem and creating another. The unlimited increase in hospitals and hospital beds had led to overprovision of hospital services and increase in cost for hospital services, which then led to reforms (e.g., prospective payment for hospital services) aiming at control of hospital costs. Grant incentives for medical professionals to work in underserved areas

Grant incentives for medical professionals to work in underserved areas

Governments often provide grants for the education of medical professionals under a contract whereby the trainees are obligated to work in government-designated areas for a number of years after their graduation. This incentive was used as a strategy in China and most former communist countries. However, there are few publications available for these countries. Here we provide an example from the United States, the United States Emergency Health Personnel Act of 1970. According to this Act, the National Health Services Corps was established, which provides up to four years of medical education support including stipend, tuition, and other educational expenses to students of medicine, osteopathy and other health professions in return for up to four years of clinical medical practice to civilian populations in health manpower shortage areas in the United States. Some 15 000 students participated in the programme and the programme provided 42 000 clinician-years of medical services to the underserved areas from 1973 to 1999 (43). This programme successfully improved geographical distribution of physicians and increased access to primary care for the underserved population (44). Policies such as this can increase general physician numbers and their primary care services in underserved areas. This can not only improve access to care by the underserved population, but also improve the mix of health interventions of the nation. However, little research on the change in the mix of health interventions is available. Funding the production of cost-effective products

Funding the production of cost-effective products

Some cost-effective medical products involve high production costs and low profit returns, and private producers are usually not willing to produce them. In order to have these cost-effective products available, governments may provide funding to promote their production. Two typical examples are the production of penicillin and the production of vaccines. During the Second World War, there was a great need for penicillin, but production was on a small scale. The United States government made an enormous financial contribution, allowing expansion from laboratory to pilot plant, and finally to mass production in an amazingly short period of time. The end result was an enormously successful postwar antibiotics industry. In many ways, this new industry was a product of the government's prominent role directing pharmaceutical research, development, and production (45).

Production of vaccine usually needs government support for the following reasons: the demand is relatively fixed, cost of production is high because of the small production scale, and private industries are reluctant to produce vaccines. But since vaccine is a cost-effective product, both low-income and high-income countries have subsidized private industries for production. Failure to provide government support would lead to a reduction in vaccine production. For example, in the United States there were 15 producers in 1965 and by 1981 there were only two left. In Canada, the existing two producers in 1981 were constantly threatened because of lack of proper recognition of the essential character of their operations and the high costs and risks. It has been suggested that government subsidies should be made available for up to 50% of the equipment and construction costs of the facilities required to produce biologicals; that the costs of expensive and unprofitable biologicals that are nonetheless essential for health protection should be partially paid by government; and that government should underwrite the stockpiling of reserves of biologicals required for emergency situations (46).

Funding the production of cost-effective medical products, particularly the products with positive externalities, seems to be an unshakeable role of the government. In countries where such products are produced by publicly owned entities, the assurance of government budget is important for the production of enough to meet the need of the population. In countries where such products are produced by private entities, governments should provide subsidies. Profit-oriented private producers tend to direct their production to the products that make the highest profit and have a potential of market expansion. Private producers are not willing to provide products for which the demand is relatively low, and which can generate little or no profit, even if they are cost-effective. In these cases, government financial support is of key importance. Theoretically, the government policies that improve the availability of such products will improve allocative efficiency of health interventions — not only increasing the quality of care and the level of health of the people, but also increasing the use of these products and improving the mix of interventions. Nevertheless, we did not find documented evidence of this last effect.

Managing capital and technology

Capital and technology play an important role in the health care system and represent a major part of health care cost. Although little concern is paid to the effect of capital control on the mix of health interventions, it may have an effect on this mix because some capital investments are directly related to specific services (e.g., hospital beds are related to hospital admissions, and computerized tomography (CT) scanners are directly related to CT examinations). Limitation on these capital investments would mean the reduction of these health interventions relative to those that are not under control. Pharmaceutical products represent a considerable proportion of total health cost, and the limitation of drugs which have more cost-effective alternatives will have effects on allocative efficiency of heath interventions. The specific tools that have been used for these purposes include Certificate-of-Need, capital expenditure cap, a capital moratorium, and technology regulation.

Certificate-of-Need

Certificate-of-Need (CON) is a US policy for containing health care costs through the control of capital investment in a private-dominated medical market. There might be similar mechanisms in other countries, but sources are hardly available. CON is a state (not national) regulatory mechanism for review and approval by health planning agencies for new construction or expansion of health care facilities, including nursing homes, the expenditure on which exceeds US$250 000–600 000. The ratification of a capital investment proposal is based on an evaluation of the need of the community an applicant serves. Other requirements may include the willingness to provide care to the poor and uninsured, and demonstration of quality of care of the applicants. This programme, funded by the federal government, was established with the passage of the National Health Planning and Resources Development Act of 1974. Its objective has changed over the past two decades. Originally it aimed at restructuring health care delivery by controlling the distribution of facilities and capital; and later it has been used to contain health care costs. By the end of the 1970s, most states in the United States had adopted the CON policy (47).

The effectiveness of CON has been debated over the years and strong disagreement persists over whether it has in fact reduced expenditures and encouraged more efficient use of resources. CON is one of the regulatory programmes that has come under the greatest scrutiny, and an extensive literature on CON performance has developed. One of the earliest studies, using a before-and-after design and based on a nationwide survey, found that CON legislation did not significantly lower hospital investment (48). Another study, which has received the widest recognition, used cross-sectional analysis of aggregated data to evaluate the effect on limiting hospital investment (49). Their conclusion was that CON had done little to reduce the level of hospital capital expenditures. After correcting for the shortcomings of the former studies (a short period, which does not allow the lag effect of CON to be considered; the neglect of the evolution of the programme over time), Howell did not find any effect of CON on hospital investment (50). Later studies showed that CON programmes had failed to control costs and had little impact on access to heath care for either the poor or geographically underserved regions(51). Generally, according to several literature reviews, the CON programme failed to control overall health care costs; served little to improve the distribution of health care facilities and equipment; and there was no evidence that the programme increased access to care. The proponents believed that CON had only added to health costs by bureaucratizing the planning process and obstructing the development of integrated delivery systems (52).

There is some evidence, however, that the programme had an effect on the containment of nursing home beds. Harrington et al. conducted an impressive study on the effect of CON policy on change in the number of nursing home beds in the United States. This study covered 49 states and the District of Columbia over a period of 14 years, from 1979 to 1993. Controlling for all other variables in the model, they found that the CON policy had a dampening effect on the growth of the number of nursing home beds (53). Some other researchers found that CON had been successful in limiting the supply of nursing home beds and it was a factor in lowering state Medicaid expenditure on nursing homes (53,54) Regulators argue that if the programmes did not meet their goals, it was because of an absence of clearly defined state health planning criteria, not CON per se.

Partly because of the failure of CON to control cost and partly as a consequence of the entry of the Reagan administration with an anti-regulatory platform and a strong interest in using market incentives rather than regulatory controls to contain costs, by the middle of the 1980s, federal financial support for the programme dropped sharply and the federal requirements for CON were removed. Many states abolished the CON programme, and other states modified it by increasing the control threshold and cutting back the number of projects reviewed (55). However, 40 states retained their CON requirement for nursing homes in 1991 (56), probably because of its effect on the number of nursing home beds.

In general, the CON programme in the United States is not a success story. The reasons for failure are multiple. First, CON agencies evaluate the need for each submitted proposal independently, with little regard for the total cost of all projects combined. Indeed, many observers believe that the failure of CON programmes to control costs is largely a result of this lack of competition for a limited pool of resources (57). Under an open-ended CON review process, an unlimited number of projects could be approved if applicants could demonstrate that the proposed services were needed. Second, CON programmes generally have played highly passive roles, awaiting hospital proposals instead of actively engaging in an effort to manage hospital capital investment decisions. Health planning agencies even at the peak of their popularity had little ability to ensure that their plans were implemented (57). Third, the process is subjective, and there is little agreement on the specific criteria for decision-making. A review of CON programmes in 20 states found a variety of different provisions with regard to standards to be applied by the reviewing agencies (58). Fourth, the review process was politicized. The allocation of sophisticated medical technologies depended much on lobbying by hospitals, communities and equipment suppliers (59-61). As a result, there was a high approval rate of requests. Even if a proposal was rejected, there would be a high chance of approval after modification according to the requirement of the CON regulators (62,63) Last, the project applicants often took advantage of loopholes: some heath care providers tried to evade CON by putting their expensive medical equipment in non-institutional settings. In the 1970s, this tactic led to CT scanners in physician offices, and in the 1980s magnetic resonance imaging (MRI) devices in the same settings (47).

Capital cap

A capital cap in the United States is a regulatory tool to limit the total capital investment for a period of time (usually a year) in a specified region. Capital cap is often used in conjunction with CON, with the former defining an upper limit of total capital expenditure of a region, and the latter determining the distribution of capital within the region. In other words, while capital cap is used to control the total amount, CON can be used to adjust the geographical distribution and capital structure. A capital cap provides an explicit ceiling for the sum of all CON-approved capital expenditures, but individual applications are judged on their own merits relative to other proposals. Among the factors considered are relative public need, relative cost and relative severity of the need for modernization or replacement (57).

Capital cap plans were imposed in many states in the United States (Kentucky, Maine, Massachusetts, Michigan, New York and Rhode Island). For example, Kentucky set its 1984 cap at US$ 166 million to be adjusted by the construction cost index for the two succeeding years. This was a substantial reduction from the US$ 369 million in capital expenditure approved in 1982. In Rhode Island, a state limit was imposed on capital expenditures, and affordability as well as community need were included in CON reviews (64).

With capital cap, the cost control function of CON can be exerted more effectively. Unless the capital allocation process is to be "first come, first served", the existence of a cap forces the CON agency to prioritize programmes. Aside from political considerations, a CON agency presumably would choose those projects that would be most beneficial to the community, rather than simply approving all reasonable projects. Capital cap has been effective in control of capital investment and growth of health care costs (57). However, there is no documented evidence of effects on quality, health, utilization and the mix of health interventions. The demise of the Health System Agencies and the perceived failure of the CON programmes in many states, combined with the free market spirit of the Reagan administration, resulted in a decline in interest in capital caps.

Capital moratorium

Capital moratorium is a policy that imposes a temporary ban on capital investment. As of 1983, a total of 18 states in the United States had implemented such moratoria, most of which applied to nursing home beds, although some extended to hospitals (57). For example, New York imposed a one-year ban on all hospital construction in 1983 (65). A moratorium on free-standing birthing centres was imposed in Georgia. Capital moratorium is an urgent and strict measure to control inappropriate capital investment. Until the ban is removed, all specified investments are prohibited. However, the ban is usually short-term; and if there is no effective capital control policy after removal of the ban, there is usually a rapid growth in capital investment (66). There is no evidence of its effect on quality, health, utilization and the mix of health interventions.

Technology regulation

Technology regulation is concerned with providing information on the efficacy, safety and cost-effectiveness of specific technologies and with regulating what technologies are allowed into health care markets. Health care technology is defined as the drugs, devices, and medical and surgical procedures used in health care, and the organizational and supportive systems within which such care is provided (67). Medical technologies that have purchasing or operating costs are commonly labelled 'big ticket technologies' (68). In the interest of cost containment, the big-ticket technologies have been paid major attention over the last two decades.

Most health economists believe that the increasing development, adoption, and use of new medical technologies accounts for a major part of the rise in health care expenditures (69), and that the primary determinant of growth in real medical spending has been the adoption and diffusion of new technology (70). It has been estimated that technology, using a broad definition, accounts for as much as 50% of the increased cost of hospital care (71). Even without rising costs, technology regulation is necessary because of concerns with the effectiveness and efficiency of health care provision. Choices among technologies have to be made at different levels of health care systems (72). Some choices are made at the national or regional policy level, as when laws and regulations prevent the purchase of equipment or the provision of certain services that are costly and have uncertain effect on health outcomes. Some choices, however, are at the operational level of clinical practice and are made by hospital administrators, heads of clinical departments, and health care providers. All these choices are apparently directly related to the mix of health interventions. International experience shows that the most cost-effective mix of technology is often not provided (67) because:

- Technologies are accepted for general use without evaluation.
- Technologies are accepted for use before evaluations are completed, making it extremely difficult to act upon subsequent results which suggest the technology is not of benefit.
- Technologies are over-supplied.
- Technologies are used for conditions beyond those covered by the evaluation.
- Technologies are used when an equally effective and lower-cost technology is available.

The focus here is the choices at system level, namely information on the effectiveness and efficiency of new technology and regulation of entry of technology to the health care market. Policy tools related to the quantity and geographical distribution of technology are captured by regional planning (discussed above); and tools related to the use of technology will be addressed under the headings of informing choice for providers, and economic incentives for providers. There are two interrelated policy tools related to technology regulation: technology assessment and technology control.

Technology assessment refers to a governmental mandate that requires medical technologies to be evaluated and information related to efficacy, safety and cost-effectiveness provided to assist the decision on the adoption and use of these technologies. Evaluation is usually based on randomized control trials and economic evaluation methods. This is the first step of technology regulation and also the prerequisite for technology control. Technology assessment is important because the lack of such information will lead to blind introduction of new technology, which is not effective and may even be harmful. For example, Beeson (73) compared treatments recommended in a 1927 textbook of medicine to those recommended in 1975. He considered that 60% of the remedies in 1927 were harmful, dubious, or merely symptomatic; only 3% provided fully effective treatment or prevention. The information provided through technology assessment is important for the mix of health interventions offered because it affects both government permission for a technology entering into the market, as well as the choice of health interventions by providers.

The effectiveness of technology assessment depends on what information is provided and its quality. In some high-income countries, such as Japan, pharmaceutical industries are required to submit information on efficacy and cost-effectiveness of a new product before it is licensed, but the information is provided by the industries themselves and the quality and validity of assessment are often questionable (74). Trial design is often restricted and defective. For example, as Maynard & Bloor stated, until recently trials in cancer treatments tended to focus primarily on the effect of the intervention on the size of the tumour and survival duration; this focus ignored both the quality of life during survival and the costs of treatment (75). In most low- and middle-income countries, valid assessment of new products is hardly undertaken. Most countries require evidence on efficacy and safety for licensing new drugs, but rarely require evidence of cost-effectiveness (76). Studies have been mounted to examine the effectiveness of major equipment, such as CT and MRI, but only rarely on the effects on health outcomes and on the costs (9). The poor design of trials, and incomplete data, provide little information in making choices that help the improvement of allocative efficiency of health resources. Given information on efficacy or effectiveness, without consideration of cost, the adoption and use of a new technology can improve health outcome, but not necessarily in a cost-effectiveness manner. Technology assessment has been promoted and even regulated by the governments in OECD countries, but none of those countries has regulated the information that must be

provided by technology producers that is necessary to enable governments and purchasers to choose cost-effective health interventions (9).

Technology control is a policy tool to regulate and restrict undesirable technology entering the health care market. Effectiveness of technology control depends, in part, on the completeness and validity of information derived from technology assessment. However, the information provided is meaningless unless it is used by different parties to make choices among various health interventions. Here, the major concern is whether government can use this information in issuing licences for a technology to enter the market.

There is substantial experience that governments have failed to control the adoption of undesirable technologies as a result either of lack of information or of failure in information use. In most countries, criteria for licensing pharmaceuticals are scientifically proven efficacy and safety. These include results of phase I and phase III studies. However, in many countries, only a marginal beneficial effect of the new drug demonstrated with a small sample is sufficient to fulfil the efficacy criteria, and cost-effectiveness is regarded as of no importance. This has led to more and more admissions of merely minor modifications of active substances, instead of real product innovations. For example, among the drugs on the market in Germany, only one-third are of proven efficacy (77). In Canada, from January 1991 to December 1995, a total of 404 new patented drug products were marketed for human use. Of that number, only 33, or just over 8%, were thought to be either breakthrough medications or substantial improvements over existing therapies; the rest were line extensions (43%) or offered moderate, little, or no therapeutic improvement (49%) (78). High-tech equipment is often introduced into the market before scientific assessment. For example, few studies are available to demonstrate more than a marginal superiority of linear accelerator against cobalt teletherapy in treatment of most malignant tumours; there was little evidence of cost-effectiveness before the Extracorporeal Shock Wave Lithotriptor (ESWL) was introduced as compared with percutaneous nephrolithotomy (PNL) and with surgical removal of kidney stones (44). As another example, electronic fetal monitoring devices, which are considered not to be effective in many high-income countries, are diffusing into low- and middle-income countries (79).

Chapter 3. Economic Incentives for Providers

Health care providers are economic beings. Some argue that doctors are utility maximizers (80-82); some argue that they are income maximizers (83,84) and others argue that a doctor's behaviour is driven by his or her income target (85-86). Some researchers state that hospitals are profit maximizers (87,88); some say that they are utility maximizers (85;89-91); and others insist that hospital behaviour is driven by quantity maximization (92-96). Although there seems to be no consensus on the objective function of health care providers, all the authors cited believe that money is an important argument. Because there is asymmetry of information between providers and users, and also because of the incompatibility between doctors' and patients' utility functions, the provider can hardly be a perfect agent for the patient. Driven by their economic motives, the providers may behave in their own best interest, to the detriment of the patient, at the expense of health system performance. However, money can be used to alter providers' behaviour towards the interests of patients. These are the theoretical reasons why payment systems to health care providers become important.

The payment system is concerned with how and how much the providers of health care are paid. The mode of payment can create powerful incentives affecting the provider's behaviour, and the changes in behaviour will affect the quantity, quality, costs and efficiency of health interventions. While it is true that some current payment systems appeared as early as the time when modern medicine came into existence in the 19th century, the development of new payment methods was largely driven by the rapid increase in the costs of social health insurance schemes all over the world in the second half of the 20th century. Presumably, payment mechanisms can affect a provider's choice of alternative medical recommendations to various patients, and hence the allocative efficiency of health interventions.

We address each of the available payment methods and explore their possibilities as a policy tool. They will be discussed in the following order: fee-for-service (FFS), capitation, salary, daily payment, case payment, line budget, global budget, and performance-related pay. The payment methods are not categorized as to whether they are used for paying doctors or hospitals because some of them (e.g., FFS, capitation, and global budget) can be used to pay both. Because the level of payment, payment structure, motivation for prescribing, and the degree of payment integration are also important dimensions of payment systems, they will be discussed independently under the headings of rate setting, fee structuring, separating dispensing from prescribing, and payer integration. Managed care is regarded as a combination of various tools, so it is the subject of a separate section.

Fee-for-service

In fee-for-service, health care providers are reimbursed based on specific items provided, such as doctor consultations, X-ray tests, surgical operations, and so on. In broad terms, FFS also includes itemized charges for medical products and drugs, because medical labour services are often provided in company with material products. FFS is similar to traditional piece-rate payment in that the payment is provided according to the number of pieces (items) for a specific type of work (service or product). As with piecework in the earliest days of employment,

FFS payment for medical services began as soon as professional medical practice came into existence (97).

FFS payment can be further divided into three subgroups, namely open-ended fees, negotiated fee schedules, and regulated fee schedules (98). The traditional type is an open-ended fee charged by the doctor according to the market. This type of payment prevailed in the medical market when medical care was less organized, regulated and planned. Although this type of payment has been declining since the early days of this century, it is still popular in some countries such as Canada, China, the Republic of Korea and the United States (under indemnity plans). Private practitioners in some countries also charge a market rate, as in the United Kingdom. The negotiated fee schedule came into existence along with the establishment of health insurance schemes. To reduce the cost of services, third parties (social health insurance schemes or private health insurance companies) often negotiate with providers or their associations a set of standard charges for the items of services. This system exists in Belgium, France and Germany (99). Canada and the United States increasingly use the negotiated fee schedule for their social health insurance programmes, and managed care organizations use it for both preventive and curative services in combination with capitation payment. A regulated fee schedule is a set of fee standards regulated by the government. This practice exists in many countries, such as China and Japan.

FFS is a traditional method of paying office-based doctors by patients and third parties; it can also be used by both patients and third parties to pay hospitals and hospital-based doctors. For example, hospital services in China, Japan and the Republic of Korea are broken down into more than 2000 items; regulated fees for these items are charged to the third party if the patients are insured, or to the patients themselves if they are not insured (98). In the United States, fees are paid by patients or the third party to the vast majority of office-based doctors, working in emergency medicine, cardiology, radiology, pathology and anaesthesiology (100). Fees can be paid directly by the third party (as in France, the Netherlands and Switzerland), or by the patients and then the patients are reimbursed by the sickness funds (as in Belgium). The fees can also be paid indirectly by an agent such as the medical association in Germany, which takes the role of negotiating and managing the total budget for doctors' services, and the Social Insurance Medical Care Payment Fund in Japan (101).

Although FFS is regarded as the worst payment method because it encourages overprovision of services and drives cost up, it has some advantages. The first and most important advantage is that FFS payment reflects work actually done and efforts actually made (98). Thus it encourages the providers to work more efficiently, and hence increases their productivity. The second advantage is that scheduled fees can be used to encourage the provision of cost-effective services by setting their prices higher relative to costs, and to discourage the provision of inefficient services by setting their price relatively lower.

FFS payment has serious disadvantages that were the focus of discussion in the second half of the 20th century. The first disadvantage is that it provides strong economic incentives for the doctors to provide costly, cost-ineffective and even unnecessary services, particularly when the workload is low, the treatment options are ambiguous, and the fees are set at a profitable level (98,102). The second disadvantage is that doctors may increase the quantity of service by reducing the length of time spent on each patient or by delegating work to less qualified workers, particularly if the workload is high. Quality may suffer because of these misbehaviours by providers. The third disadvantage is the relatively high cost of administration to both providers and insurers (99). In general, FFS is recognized as favouring the internal efficiency of providers and acting against social efficiency.

Empirical studies mainly focus on testing whether FFS payment results in the overprovision of care. There is strong evidence to demonstrate the association between the increased utilization of services and FFS. One of the earliest studies showed that in the United Kingdom, where surgeons were paid either a salary or some combination of salary and capitation, the rate of surgical operations per capita was about half of that in the United States where surgeons were paid on a FFS basis (103). Several studies in the United States showed that variation in geographical rates of surgery seemed to be best explained by the number of surgeons in each geographical area paid by FFS (104). Primary care physicians who used their own imaging equipment in their offices, ordered over four times as many imaging examinations as the physicians who referred patients to radiologists for examination (105). The link between utilization rates and FFS was further demonstrated by evidence from various health systems of Western Europe and Canada, as well as the United States (106). Under FFS, doctors have frequently responded to a reduction in real earnings over time by increasing the amount of services delivered per capita (86,107-110). The comparison of medical expenditures between health maintenance organizations (HMOs), with their doctors being salaried or capitated, and providers with FFS payment showed that expenditure for the former was 10-40% less than for the latter (106). A cross-sectional analysis of over 20 000 patients who visited providers' offices during a 9-day period in 1986 showed that after adjusting for patient mix and the pattern of practice (solo or group practice), FFS payment led to 41% more hospitalizations than HMOs (111). Although the question of whether supplier-induced demand exists has not been conclusively answered because of problems of measurement, it is true that the quantity of provision under FFS payment, and medical cost are significantly higher than under other payment systems.

In spite of strong objections to FFS payment, this system has not yet been abandoned. The first reason is that doctors, who have a great deal of political power in the developed world, prefer this system because it provides autonomy for medical decisions and opportunity to earn a high income. The second reason is that some countries, such as China, Germany, Japan and the United States, have a tradition of paying the providers based on FFS. A revolutionary change may be difficult, especially in democratic countries where reforms can only be based on negotiation among different stakeholders. The third reason is that FFS has proved to work well in Canada, Germany and Japan, where it is combined with a global budget under a single-payer system. The global budget provides an expenditure cap for physician services, and the single-payer system prevents the providers from shifting costs to other payers. In implementing FFS with a global budget, the fee schedule provides a relative value or number of points for each item of service; the monetary value of each point depends on the total budget and the quantity of service provided. Because the total budget is fixed, more services mean a lower price for each item. This provides incentives for the doctors to reduce the quantity of services. This mechanism works better if the income cap of the physician is regulated as in Canada. The last reason is related to the second advantage of FFS, which is that it can be used to motivate doctors to provide cost-effective interventions. For example, the South Carolina Preferred Personal Care Plan reimburses physicians $675 for a colonoscopy in an outpatient setting and $515 for the same procedure performed in a hospital (112). Another example is that in the United Kingdom, FFS is used for paying general practitioners for their immunization and screening services (113,114), though the curative services are capitated. Following the practice in the United Kingdom, FFS is increasingly used to pay for preventive services to encourage provision of these cost-effective services.

Experience has shown that FFS can lead to more services, costlier services, and higher costs; and that the problem of FFS can be corrected, at least partially, by the global budget and single-payer system. The effect of FFS on quality and health outcome is hardly documented. Although there is no evidence whether FFS can affect the mix of interventions, it is increasingly used to pay health care providers for preventive services in countries of the European Union.

Capitation

Under capitation, the provider is paid a periodic fixed amount per insured person, and in return is responsible for the provision of a defined package of health services. Capitation payment has been implemented in Denmark, Italy, the Netherlands and the United Kingdom, and has been introduced in Costa Rica, Indonesia, and in HMOs in the United States (98). This type of payment transfers the economic risk from the third party payers to the health care providers. The provider receiving a capitated fee can be an office-based doctor or a hospital (115). The capitated office-based doctor is usually responsible for the provision of only primary care (as in Hungary) or in addition for the purchase of hospital care (as GP fundholders are, in the United Kingdom). The capitated hospital is responsible for providing inpatient care, or in addition, for providing outpatient care, such as the hospitals contracted by the social health insurance scheme in Thailand (116). The capitation fee can include both services and drugs (as with some HMOs in the United States) or can include only consultant services (98). Capitation payment may be a flat fee for each of the providers or it can be a risk-adjusted fee, based on the relative risk of the registered population. For example the capitated fee is adjusted in Germany by five variables — age, sex, whether the insuree is disabled and of working age, family size, and income (115).

The most important advantage of capitation payment is that it removes the economic incentive of overprovision, adds an incentive to provide cost-effective care including preventive services, and thus helps to control health care costs. Because the provider is responsible for providing the contracted package of services with the fixed payment, this can motivate the provider to innovate in cost-reducing technology, the use of lower-cost alternative treatment settings, and the provision of cost-effective care. This change in economic incentive is related not only to cost containment, but also to the improvement of allocative efficiency. Hornbrook predicted theoretically the effect of capitation payment on allocative efficiency (90). He divided health interventions by two dimensions — type of disease and type of interventions. The former is divided into:

- self-limiting disease — from which the patent is expected to recover completely without any medical treatment, e.g., influenza;
- acute disease — for which medical intervention is required to prevent death and restore health, e.g., appendicitis;
- chronic disease — the patient is unable to recover fully and suffers some residual disability, e.g., diabetes mellitus;
- terminal disease — which can only progress to death, regardless of medical interventions.

The type of intervention is divided into:

- prevention — that can prevent the occurrence of disease, e.g., immunization or interventions related to the avoidance of high-risk factors;

- health screening — related to the earlier detection of disease;
- diagnosis — the collection of information for assuring the type of disease and for the plan of treatment;
- emotional support — that helps to release the emotional suffering of a patient and his or her family, but is not related to the technical treatment of disease;
- therapy — the treatment of disease.

According to Hornbrook's prediction, capitation payment will encourage physicians to provide preventive services, health screening and diagnostic services because the prevention of disease can avoid treatment cost, and earlier detection and diagnosis can reduce the costs of treatment resulting from progression. However, capitation will discourage the use of cost-ineffective therapies and the provision of emotional support, for the purpose of reducing cost. He also predicted that the capitated physicians would be more willing to provide care for acute patients than for patients who have self-limiting, chronic or terminal diseases. If the change of physician's behaviour is as he predicted, allocative efficiency can be improved, only if the change in the mix of health interventions and the reduction of cost does not have negative effects on health outcome.

While capitation payment provides no incentive for unnecessary care, it may provide incentives for reducing the provision of necessary care. The first disadvantage is that the provider will try to select low-risk clients, and reject the high-risk ones if the capitation payment is not adjusted for individual risk. For example, the evidence from the United States indicates that HMOs had healthier enrollees than the rest of the population, suggesting they had selected favourable risks to some extent (117). The second disadvantage is that the provider may reduce the quality of care to reduce costs. This can be done by using low-quality premises and equipment, reduction in the number of necessary tests, decrease in the length of services, longer waiting lists resulting from too many registrations, and so on. The third disadvantage is that the patient is more likely to be referred to a specialist or a hospital than necessary, because more referral means less cost for the capitated provider. For example, the capitated payments to family physicians in Hungary covered only their own services. The referral rate was much higher than when physicians were paid salaries (115).

Several possible solutions have been suggested and tried out in the practice of capitation payment. To deal with adverse risk selection, individual risk adjustment for the capitation fee has been the focus of discussion in the past decade. Colombia, Germany and the Netherlands are just starting to use simple formulas to adjust the risk. As Barnum, Kutzin & Saxenian state, however, simple formulas may work better when benefit packages are limited; more complex formulas may be needed for comprehensive packages (115). Experience to date shows that ideal methods of risk adjustment have not yet come into existence. To assure the quality and the quantity of health services under capitation, competition is suggested in many countries and has been tried in the United Kingdom where the clients are given freedom of choice of general practitioners (GPs). The GPs have to compete for registrations by assuring reasonable quality and the necessary quantity of services. In addition, to assure quality, the number of registrations should be limited by regulation, as is in the United Kingdom. To deal with unnecessary referrals, the capitation fee should include both primary and secondary services. There are generally two different practices for the same purpose. One is GP fundholding, invented in the United Kingdom, where the GPs are responsible for providing the primary care and purchasing the defined specialist and hospital care with capitated payments. Another practice can be found in China and Thailand, where contracted hospitals are paid capitation fees by the social health

insurance schemes, and in turn they are responsible for providing both primary and secondary services. These two approaches remove the incentive for unnecessary referrals, but add the incentive to keep the patients at the primary level when referrals are needed. The latter incentive may be limited by constructing competition among providers through patients' choice.

There are substantial experiences of change in the mix of health interventions associated with capitation payment. First, studies show that the implementation of capitation payment has resulted in provision of more preventive and screening services. For example, the results of a three-year parallel, controlled clinical trial comparing a capitation system of payment for the dental care of children with fee-for-service showed that capitation offered dentists more clinical freedom. Dentists under capitation provided more preventive care, particularly advice to parents on the control of dental diseases in their children. Parents were satisfied with the preventive service their children received and were confident of their ability to control their children's dental disease (118). As another example, in the United States 80% of women aged 50–74 years enrolled in the Kaiser Permanente Plan of Northern California (in which the physicians were paid capitated fees) had received mammography screening, compared with 25% of women in this age group in the population as a whole; paediatric immunization rates were over 90% in Kaiser plans, compared with a national average of 37% (115).

Second, studies have shown that capitation payments that include drug expenses have led to more rational prescription of drugs. A before-and-after comparison with a control group showed that after FFS for pharmacies was changed to capitation fee in the Medicaid programme in Iowa, United States, the rate of generic substitution increased 19-fold (from 0.38% to 7.25%). The increase in substitution resulted in substantial saving without a decrease in health outcome (119). Moreover, it was found that capitation payment to pharmacies by the Medicaid programme had a spillover effect on the generic substitution rate of non-Medicaid patients whose costs were based on FFS. The introduction of capitation for Medicaid patients changed the rate of substitution for non-Medicaid patients from 0.67% to 3.19% (120). Following the successful experiment in one county in Iowa at the end of 1970s, capitation was expanded to 32 counties in the early 1980s (119).

Third, the mix of curative services may have changed. One study showed that capitation payment for comprehensive services resulted in 40% fewer inpatient admissions in comparison with FFS. The length of stay was significantly shorter under capitation as providers sought to minimize high-cost stays in hospitals. Length of stay had fallen from 5.9 days in 1985 to about 3.75 days in 1993 (42). GP fundholding in the United Kingdom created incentives for the GPs to provide care that was previously provided by secondary care institutions. This caused a substitution of primary care for secondary care for minor surgery and chronic disease management (121). The Cardiology Roundtable in the United States (an advisory board providing research and consulting service to the hospital industry) conducted research which involved interviews of more than 200 experts in capitation and cardiac services and predicted that diagnostic catheterization, percutaneous transluminal coronary angioplasty and coronary artery bypass grafting (CABG) surgery rates would all be reduced by approximately 60%; cardiology ambulatory care visit rates would be reduced by almost 40%; and cardiac inpatient utilization would be reduced by nearly half (122).

Capitation payment is primarily a tool to control health care costs. Studies have shown that the reduction of cost largely results from the change in the mix of health interventions by

substituting the less costly ones for ones with higher costs. However, as stated by Scott, evidence on whether these substitutions are cost-effective is scarce (121). The reduction in cost through changes in intervention mix means an improvement of cost-effectiveness only if there is no underprovision of necessary care and health outcome is not reduced or is improved. Much literature has been generated on this in the United States, where managed care is the main topic under analysis, rather than only on capitation payment. Managed care will be discussed in Section 7.

Salary

Salary pays doctors, based on the time worked. A salary can be paid on a part-time or full-time basis depending on the pattern of employment. Salary payment to doctors is quite common in planned health care systems. For example all hospital-based doctors in China and the United Kingdom are salaried. Doctors who provide care in outpatient health centres are often salaried on a full-time or part-time basis, as in Finland, Greece, India, Indonesia, Israel, Portugal, Spain, Sweden, Turkey, the former USSR and many countries in Latin America (98). Doctors can be employed and paid a salary by health funds or insurance institutions, independent hospitals, independent outpatient clinics, and other non-health organizations (such as universities and industries).

One of the most important advantages of salary is that it does not provide the economic incentive for overprovision as in the case of fee-for-service or for underprovision as in the case of capitation. The second advantage is that it could make health care planning easier, as a result of knowing the doctors' pay in advance (123). The third advantage is that the salary system encourages doctors to conduct group consultation, which is necessary for complex cases to provide the definite diagnosis and work out the appropriate treatment plan. Group consultation is difficult under fee-for-service because the work and the related payment is hard to share by a group of doctors. The fourth advantage is the lower monitoring and administrative cost compared with that under fee-for-service and capitation payments.

The disadvantages of the salary payment are stated by several authors (98,123,124). The major disadvantage is that a fixed salary does not provide incentives for doctors to work productively. Low morale is a problem, especially for those doctors with potential for higher output, because they may think they are not rewarded for their hard, and good quality work. The second disadvantage is that salary payment provides no direct incentive for doctors to recommend the most cost-effective health interventions, decrease costs, and increase health outcome. Salary payment may cause low productivity, and low quality of care, and allocative efficiency may be lower than under capitation payment. A third disadvantage is that if doctors are paid relatively low salaries, they may ask for or accept illegal payments from the patients and take kickbacks provided by pharmaceutical industries and high-tech equipment producers.

Although there are many disadvantages, salary payment is still the most popular payment method around the world. In some countries where fee-for-service payment for physician services has been popular, such as the United States, physicians are increasingly paid a salary as part of integration between physicians and hospitals, and between third parties and providers. One of the supporters of salary payment argues that it is desirable because it provides a neutral incentive to doctors behaviour, and what a doctor will recommend and prescribe depends

wholly on the need of patients, his or her medical knowledge, and the availability of resources (102). The cost-effectiveness of interventions recommended by doctors can be improved through educating doctors, providing scientific evidence, and offering medical practice guidelines. The disadvantages of low productivity, low morale, and less cost awareness can be overcome by the partial ownership of the facility by physicians, the proper design and implementation of bonus systems, and the use of various non-financial motivations.

A review of the available studies related to the effect of salary payment on the behaviour of doctors (125) included 23 papers of reasonable quality. The authors found that salary payment was associated with the lowest use of tests and referrals, compared with FFS and capitation. Salary payment was also associated with lower numbers of procedures per patients, lower throughput of patients per doctor, longer consultations, and more preventive care compared with FFS payment. These results suggest that salary payment has resulted in changes in the mix of health interventions. However, whether the changed mix of interventions is more cost-effective is not known because of lack of information of the cost-effectiveness of various interventions. The authors were not able to draw conclusions on whether salary payment has a positive effect on quality and health outcome as compared to other forms of payment systems, because only one review paper reported health outcomes, and the measure in that study was wound infection rates (which may be not a valid measure). None examined whether salaried doctors differentiated between patients on the basis of health needs.

Daily payment

Daily payment or per diem payment reimburses the institution providing inpatient services a fixed amount for each inpatient day regardless of the actual use of services, drugs and medical products. In theory it is applicable to all inpatient services including long-term care in nursing homes. In practice it is only found in hospital payment by third parties (99). This type of hospital payment is commonly used in continental Western Europe (19;126) and is being tried out in China and Indonesia in their social insurance schemes.

This type of payment provides incentives for the hospital to increase the total number of hospital days by increasing both the length of stay and the number of admissions, while reducing the intensity of care for each hospital day. Thus, technical quality may suffer as a result of insufficient services and drugs, while the perceived quality (e.g., the doctors' attitude to patients) may increase for keeping the patient longer and attracting more admissions.

This system may work well when there is a budget cap for hospital services. The low quality of care and the lengthening hospital stay can be monitored by a peer doctor, but the monitoring costs will be high. Whether this payment can reduce the cost of hospital care depends on whether the payers can effectively control the increase in the length of stay. Whether this method can improve allocative efficiency depends, in addition, on whether payers can effectively monitor the hospital to provide the necessary services and drugs.

Empirical evidence shows that per diem payment is associated with long length of stay and large numbers of admissions. For example, while the urban population in Brazil increased by 50% from 1971 to 1981, the number of hospital admissions in urban areas increased from 3 million to about 11 million, and the cost for hospital care increased dramatically. Policy-makers and researchers believed that this was driven by the per diem payment. On the basis of this

evidence, Brazil abandoned per diem payment, and case payment was adopted in 1983 to control cost and reduce the number of hospital admissions (127). Germany used per diem to pay hospitals, and that resulted in a long length of stay, 13.1 days on average in 1987 (19). The need to control costs and shorten the length of stay led to calls for reforming the payment system to hospitals, and in 1993 a law was passed to replace per diem payment by case payment (128). With the payment reform, the length of stay decreased, but the number of admissions continued to increase. As a result, the number of hospital days per 1000 population also increased. There was no evidence on whether the abolition of per diem payment led to reduction in cost and change in the mix of health interventions (129).

Case payment

Under case payment, third party payers pay a fixed amount per case regardless of the actual types and quantities of services provided. Case payment can be used both for outpatient care (such as the payment per inclusive visit that is being tried in China's social health insurance reform) and inpatient care (e.g., the Diagnosis Related Grouping (DRG) in the United States), and for both physician services and hospital care. Case payment can be a single flat rate per case regardless of the diagnosis, and can be diagnosis-based. The most popular type of case payment is the DRG payment for hospital services, which has been implemented in the United States in its Medicare programme and has been adopted or experimented with in many other countries (such as Brazil, Germany and Indonesia). The case payment method is only used by third party payers, not by individual patients.

The principle of case payment is that the costs among cases within one diagnostic group should be as similar as possible, and the needed cost and technology of administration should not be beyond the available capacity. Case groups can be as simple as only one group, and as complex as 478 groups in the United States; others may be in the middle of the two extremes, as in Indonesia (98,112). In some countries in central Asia the number of groups is huge (as many as 10 000), because the disease grouping is simply based on ICD-10, without much bundling of the diseases.[1]

DRGs represent the first and the most sophisticated method of case payment. The idea of DRG was invented by Fetter (130). It was introduced by the Federal Government of the United States as a prospective payment system for Medicare hospital patients. The grouping of patients was conducted through cross-hospital studies of the average costs for each type of diagnosis, and then the factors that affect the level of cost for the diagnoses were identified. The major factors selected are principal diagnosis, secondary diagnosis, principal procedure, secondary procedure, destination of discharge, sex, age, and length of stay. According to these factors, patients are divided into groups which reflect the difference in resource utilization for providing hospital care. In a DRG payment system, payments are made for cases according to their indices of case-weighted admissions, which are relative values reflecting the relative costs of DRG groups. The monetary value for each DRG weight, called the standardized amount, is established by considering the budget of the Medicare programme and the costs of hospital services. The adjustments of reimbursement are made with regard to areas (urban or rural hospitals), teaching responsibility (teaching or non-teaching hospitals), and outliers (length of stay and cost per DRG case that exceed the normal). In addition, the capital costs (depreciation

[1] Based on personal communications with J. Langenbrunner, World Bank, in October 2000.

of fixed assets) and the cost for direct education (the salary of interns and residents) will be passed through. Efforts have been made to incorporate these costs into the DRG rate, but because these costs vary a great deal from facility to facility, such proposals have been met with resistance from facilities with higher capital and education costs.

One of the major advantages of the case payment system is that it removes the economic incentives for the hospital to provide as many items of services as possible, as in the case of fee-for-service, and to provide hospital stay as long as possible, as in the case of daily payment. In addition, the reduction of unnecessary services will help to increase the quality of care (131,132). It is easy to operate and the administrative cost is low if the providers are reimbursed based on a single flat rate of case payment.

The predicted disadvantages are various. The first is the so-called "DRG creep", which means that the hospitals are likely to code patients into groups with a higher point value (or index) for more reimbursement. The second disadvantage is called cost shifting, which means that, while the costs for DRG patients are controlled, costs will be shifted to the non-DRG patients, and as a result, the total cost to society does not fall. Another type of cost shifting occurs when the provider increases the quantity of pre-admission tests and conducts premature discharge. As a result, the costs are shifted to outpatient services, home service care and nursing home care. Still another type of cost shift is that the provider may skim the costly cases and shift the costs to other providers. As Omenn & Conrad stated, those cases with the highest level of reimbursement relative to their costs will be sought after by the hospitals (133). The third disadvantage is that this payment system provides the incentive for the provider to undertake unnecessary admission and readmission. This may increase the total cost of health care, and decrease the quality of care because of the interruption of care. The fourth disadvantage is the likelihood of quality reduction caused by the reduction in necessary care, including the length of stay, services and drugs. It is clear that an effective implementation of case payment depends on whether the payment policies can ensure that the cases are assigned to the right diagnosis groups, the patients are not needlessly transferred from one provider to another, and a certain level of quality is maintained (99).

Empirical studies generally showed that case payment can reduce the length of hospital stay, reduce the daily cost due to reduction in use of optional technologies, and increase the number of admissions. One of the earliest studies, in New Jersey, showed that compared with hospitals with cost-based payment, the hospitals with DRG payment had a length of stay 6.5% shorter; the cost per admission was 14.1% less; and the number of admission was 11.7% greater(134). A before-and-after comparative study showed that the length of stay fell about 9% after the payment system changed from cost-based payment to DRG (135). The RAND before-and-after comparative study, which involved five states in the United States, showed that the implementation of DRG reduced hospital length of stay by 24% from 1982 to 1986(136). In a controlled before-and-after study aimed at assessing the effect of Medicare prospective payment on the use of medical technologies in hospitals, Sloan et al. (137) found that the use of many non-surgical procedures and routine tests declined, because under DRG, physicians are usually asked by hospital administrators to order fewer laboratory tests and examinations (138). A very impressive national analysis showed that: during the first year of the implementation of DRG in 1984, the average length of stay for Medicare patients dropped from 9.33 to 7.69 days, and the growth rate of DRG payment per case dropped from 10% in 1984 to 4% in 1987 (139). As stated by Rosko & Broyles, it is clear that the implementation of the DRG payment can shorten the length of stay and reduce daily costs as well as the cost per admission (134).

It is generally controversial whether case payment can reduce costs, because it is observed that: there is an increase in the number of admissions; some of the costs of patients are shifted from inpatient to outpatient care because of 'quicker and sicker' discharges; and some of the costs are shifted to non-case payment patients. Coulam & Gaumer concluded that the change from cost-based reimbursement to case-based reimbursement led to a substantial decrease in the growth rate of the cost of the United States Medicare programme (140). Another study showed that the implementation of DRG was associated with a large reduction in the rate of hospital cost inflation (141).

The reduction of Medicare cost and hospital cost does not necessarily mean the reduction of overall health care cost for society. One of the challenges is that the decrease in cost resulting from a decrease in length of stay can be offset by an increase in the number of admissions. For example, in Germany, the switch of per diem payment to DRG in 1993 led to a decrease in the length of stay but an increase in the number of admissions; as a result the rate of cost increase did not change much (128). In Brazil, cost escalation led to case payment for hospital services in 1983, but experience showed that case payment did not reduce hospital admission and hospital cost (127).

Contrary to expectations, however, during the years of DRG implementation in the United States, the number of admissions actually dropped (from 11 546 thousand in 1984 to 10 722 thousand in 1988), and the widespread fear that DRG payment would cause hospitals to increase admissions proved unfounded (139).

Another challenge is the possibility of DRG creep. The analysis of national data by Chulis showed that there was a continuous increase in the case-mix index from 1984, suggesting the possibility of "upcoding" or DRG creep; but this was not conclusive because other factors were not ruled out (139).

Still another challenge is the likelihood of cost shifting from inpatient to outpatient care and nursing home care. Miller et al. found that along with the implementation of DRGs in the United States, patients were shifted from inpatient to outpatient services of the hospitals (because of earlier discharges), and the cost for outpatient hospital services increased at an annual rate of 15% from 1990 to 1994 (142). Analysis of national data showed that while the rate of increase in inpatient Medicare expenditure during the five years before DRG payment was 45% (adjusted for inflation) and 3% during the five years after DRG payment, the rate of increase in expenditure for outpatient services was 64% before DRG and 74% after DRG (139). It was generally believed that case payment would induce shorter hospitals stays, but there was a significant increase in the number of patients with disabilities and health problems admitted to long-term care facilities (143) or sub-acute settings (144) upon hospital discharge. As a result the cost saved from the shortening of length of stay was offset by the increase in outpatient and nursing home costs.

Another unintended consequence of case payment in the United States was costshifting to non-case payment patients. Traditionally, hospitals and other providers used revenues from insured patients to cross-subsidize uninsured patients and patients for whom payments were below marginal cost. With case payment, hospitals' profit margin from Medicare patients is reduced, and hospitals are thus less willing to provide care to the uninsured poor. Studies showed that cross-subsidization was no longer effective for covering the cost of care for the uninsured (145,146).

Another concern related to case payment is whether it can affect quality of services and health outcome. Along with the increased use of DRGs comes the fear that some of the resulting reductions in length of hospital stay and costs will be achieved at the expense of lowering quality of care. While there is an argument that hospitals with DRG payment are expected to adopt control systems that are explicitly aimed at enhancing both efficiency and quality (147), evidence related to the impact of case payment on quality is limited and inconsistent. A before-and-after comparative study of the RAND Corporation analysed the effects of Medicare prospective payment on patterns of hip fracture care (148). Although the mean length of hospitalization fell from 16.6 to 10.3 days and the number of physical therapy sessions also decreased from 9.7 to 4.9, the proportion of patients discharged to nursing home care increased from 21% to 48%. After six months from the day of hospitalization, 39% of patients remained in nursing homes post-DRG as opposed to 13% pre-DRG. According to the authors, these results suggest deteriorating quality of care and health outcome, and overall cost increase. Another study showed that DRG payment for non-insulin-dependent diabetics hospitalized for glycaemic control resulted in the realization of shorter hospital stays and fewer tests, but led to adverse effects in terms of less patient education, more emergency room visits, more hospital admissions and worse glycaemic control (149).

In contrast to these studies, some research found that the shortening of hospital stay and the reduction in use of technology improved or did not affect quality. In an impressive analysis of the effects of prospective reimbursement, Kahn et al. assessed the quality of clinical practice and health outcomes before and after the introduction of DRGs (150). Quality of care was measured by two dimensions: process (explicit process criteria) and health outcome (stability of condition on discharge, mortality at 30 days and 180 days after admission). The findings were that process quality was not adversely affected; there was no difference in 30 and 180 days mortality; and the health condition of DRG patients was less stable on discharge. In another study, Coulam & Gaumer found no evidence that the reduction in the number of diagnostic tests and therapeutic activities per case was associated with negative consequences for the quality of care (140). A literature review by Wiley suggested that the implementation of DRGs did not negatively affect the quality of care for Medicare patients, because studies in the United States found that quality of care for Medicare patients had been improving during the period since the introduction of DRGs (151). The United States national data showed that although the number of deaths per 1000 admissions increased 3.7% during the first year of DRG payment, the number of deaths per 1000 Medicare beneficiaries decreased by 4.1%; the former can partly be explained by the decrease in number of admissions with less severe conditions and partly by the increase in the severity of admitted inpatients; and it is unlikely that DRG payment had a negative impact on the mortality rate (139).

DRGs have been used beyond the United States, and experiences with these uses are quite positive. In the Medicare programme, the hospital receives a fixed DRG-based amount for the institutional services associated with each admission; individual physicians, including surgeons, bill Medicare on a fee-for-service basis for their services provided within hospitals. In a Medicare test, compensation for professional services was pooled together with the institutional payment into a single fixed dollar amount. Thus reimbursement was based on a per admission rate for all services including drugs, medical products, institutional and professional services. The heart bypass demonstration is perhaps one of the most successful payment demonstrations that Medicare has ever undertaken (152). The demonstration showed that providers' response to the inclusive DRG payment incentives was unexpectedly large. During the first three years in the management of surgery for coronary artery bypass grafts in patients paid by DRG, it was found

that there were less operating room time and quicker turnovers; more generic drug substitution; reduced length of stay in the surgical intensive care unit; expanded use of clinical nurse specialists for patient management; fewer medical consultations; and earlier discharge from the hospital. This demonstration showed that there was a substantial cost saving, and the evaluators detected no evidence of adverse impacts on either the inpatient or one-year mortality rates.

The United States DRG system has been modified and used by many European countries and Australia as a way of financing public hospitals under a global budget (151). In these countries, DRGs are not used on a case-to-case basis to pay hospitals, but rather they use DRGs to measure the case-mix of inpatients and finance hospitals based on their case-mixes. The primary reason for these countries to adopt DRGs is not to control cost, as is the case in the United States, but to improve equity of hospital financing and promote hospital productivity, because it is recognized that hospital financing based on the number of beds lacks motivation for the hospital to eliminate a long waiting list, and hospital financing based on the number of cases treated is inequitable because of the differences in their case mix. The use of case-adjusted financing for hospitals is expected to create no problem of cost overrun because the total cost of hospital care is capped under the global budget: to increase the number of admissions and hospital productivity; to reduce the use of technologies (tests, procedures and drugs); and to improve allocative efficiency if the quality and health outcome of hospital care can be maintained or improved. The available evidence at the moment is reported by Duckett (153). In Australia, DRG payment was used for case-mix funding combined with global budget, starting from 1993. The objectives of the programme were to reduce hospital expenditure and to reduce patient waiting lists. After six months, evaluation showed that the two objectives were achieved. Total expenditure on hospitals was about 5% less in 1993 compared to 1992. However, the number of patients treated in the period July–December1993 was about 5% higher than the number treated in the same period in 1992. This study did not specify the changes in health outcome of the patients. The lack of information on health outcome constitutes a major problem in evaluating the efficiency of case payment. A review of literature by Donaldson & Magnussen concludes that there is no evidence regarding the effect of DRGs on efficiency (126). They argue that the use of DRGs within a global budgeting system has more potential, but only once efficient clinical practice has been established and hospitals are monitored to improve the quality of their services; that improvements and maintenance of health outcome, and not the number of patients treated, is the appropriate product of hospital care; and without much data on patient outcomes, it will never be possible to determine which funding mechanism will promote efficiency.

Line budget

With a line budget, a health care provider is paid an amount per period (usually per year) for defined responsibilities of service provision; the total amount is broken down into several items, such as salaries, drugs, equipment, maintenance and the like; and regulations prohibit the managers from switching the funds among the line items, unless approval is provided by the funding agency. A line budget can be either variable or fixed. The former means that the budget is only a reference value of spending for the budget period, and it can be further divided into the open-ended budget and the target budget. An open-ended budget means that if the budget for this year is overrun, the payer will provide additional funds and the budget

for the next year will be increased accordingly; and if there is a surplus by the end of the year, the surplus will become a part of next year's budget. The target budget means that if the budget for this year is overrun, the payer will provide additional funds, but the budget for the next year will be less than it would otherwise be, as a penalty for the overrun (101). If there is a surplus, it can be retained by the hospital as a bonus. The fixed budget is also called a budget cap or budget ceiling, which means that after the amount of budget is decided it will not be changed, and providers are responsible for assuming all profits and losses.

A line budget can be provided by governments to their affiliated health facilities; by third parties to their own health institutions; and by third parties to independent providers. It is commonly observed in public facilities directly financed by government finances or indirectly financed by government through insurance funds.

The open-ended line budget, which is increasingly being abandoned in the world, provides an incentive for inefficient use of resources and rapid spending towards the end of the budget year, since it provides a blank cheque for a hospital to spend the money without any economic risk. The line budget is so rigid in the spending pattern that it allows no choice of the best and the least costly combination of inputs to produce services. These theoretical predictions and the widely observed (but rarely scientifically investigated and reported) inefficiency in using the budget probably are the major reasons for the abolition of the open-ended line budget. Target and fixed line budgets provide an incentive for providers to use the budget carefully, but they allow no transfer between line items and thus allow no choice of the least costly combination of inputs. As autonomy is granted to public hospitals, the line budget (regardless whether variable or fixed) is increasingly being abandoned. There is no evidence on how to manipulate the budget items to improve productive efficiency of hospitals and the mix of health interventions.

The slight available literature takes it for granted that the line budget is a payment method for hospitals. The line budget is often also used by governments to allocate resources among different types of public facilities and public health activities. When the government provides a budget for their public facilities, it is quite usual to specify where the budget goes. The government budget can be itemized according to the types of services (immunization, water and sanitation, maternal and child health, etc.); the level of facilities (tertiary hospitals, secondary hospitals and primary health centres, etc.); types of health facilities (e.g., health promotion and disease prevention, curative health services, and rehabilitation); and the types of major inputs (e.g., equipment and buildings). Theoretically, allocative efficiency can be improved by manipulating the government budget lines to increase the provision of more cost-effective health interventions and decrease the provision of less cost-effective health interventions. But there seems to be no reported evidence on how the government budget should be manipulated and to what extent the manipulation of government budget affects allocative efficiency.

Global budget

Under a global budget, providers are paid a fixed amount for a given period of time for providing specified services, and can make discretionary use of the budget. While the concept is simple, the types of global budget vary depending on the flexibility of budget, the types of

providers, the number of providers, the number of payers, the budget cap, and the basis of determining the budget.

According to the degree of flexibility, the global budget can be divided into two types — soft and hard. A soft budget means that if there is an overrun the purchaser assumes part of the financial responsibility; a hard budget means that the budget is fixed, and if there is an overrun the purchaser does not assume any responsibility, and all the financial risk is shifted to providers. Ordinarily, a global budget refers to a fixed budget. Global budgets can be used for hospital services, physician services, pharmaceuticals, and for both services and drugs. A global budget can apply to each of the individual providers (each hospital is provided a fixed budget) and all providers (the budget for each hospital is not fixed, but the total budget for all hospitals is fixed). According to the number of payers, global budgets can be classified as a single-payer budget which constitutes the only sources of provider financing, or a multi-payer budget which is only a part of a provider's revenue. According to the target of budget cap or what are capped, global budgets can be divided into those that set an upper limit on the total revenue of providers, and those limiting the total spending of providers. The type of global budget also varies depending on the basis on which the budget is set. The alternative bases include historical spending and activities (block contract), the number of staff, the number of beds, the quantity of service provided, and provider performance.

Global budgets for hospitals have been implemented in Canada, the United Kingdom, and all countries with national health services; a global budget for physician services has been undertaken in Germany, Canada, the Netherlands, and the United Kingdom; and only one county (Belgium) has issued a global budget for pharmaceuticals (154). Available evidence refers mainly to global budgets for physician services in Canada and Germany, global budgets for hospitals in most OECD countries, and hospital spending caps in the United States.

Canada and Germany follow a similar pattern of global budgets for physician services: there is a total budget cap for all physicians by a single payer, which is central to containing aggregate expenditures for health (155). Physicians are paid on a FFS basis, where the level of fee for a specific service depends on the amount of fixed budget and the quantity of service provided; the more service provided, the lower the fee for a service, with total spending for physician services unchanged. In Germany, the global budget for physician services is based on negotiations between sickness funds and the medical association (19). In Canada, the global budget is based on the negotiation between the government and the medical association. This system is very powerful for controlling the cost of physician services because the size of the total budget is fixed regardless of the quantity of service actually provided, and all financial risks are shifted to providers. Physician services are treated like common-property resources, characterized by jointness in use, indivisibility, subtractability, and excludability (156). According to theory, under FFS and global budget, there will be an overprovision by physicians because the more a physician provides the more he or she will get from the fixed sum. But whether this will happen depends on a number of factors: the degree of financial pressure on physicians (the greater the pressure, the more likely is overprovision); participation in decision-making by physicians; and rules, monitoring and sanctions (157). This payment system provides no incentive for the providers to improve or maintain their service quality. Without monitoring and related sanctions, the quality of care is expected to suffer.

In Australia and many European countries, the integration of case-mix-adjusted hospital financing with a global hospital budget constitutes the major form of hospital payment

(62,154,158). According to this system, a hospital is paid based on the product of the number of admissions and the case-mix index. The greater the number of admissions and more severe the case treated, the more payment the hospital can get; but the distribution of the budget among hospitals is within the limit of the budget cap. The incentive provided by this system is similar to case payment as discussed earlier, but a budget cap is expected to be a powerful tool to control hospital cost.

An expenditure cap for all hospitals in a region is another type of global budget. Since 1974, hospitals in Rhode Island in the United States have participated in annual negotiations with state officials and insurers to determine the allowed increase in statewide hospital costs (the Maxcap) for the next fiscal year. Individual hospital budgets may be above or below the quota, as long as the total increase in cost for all hospitals in the state does not exceed the negotiated amount (159). Another example in the United States is Vermont which was reported to plan to enact legislation that would make it the only state to implement a global health budget, capping both public and private medical spending (155). The effectiveness of this type of global budget on cost control depends on the penalty if spending exceeds the regulated cap, and it seems that hospitals have no incentive to reduce expenditure if there is no penalty.

In general, the global budget has proven to be a popular cost-containment technique, especially in the wake of the large cost increases most countries experienced in the 1970s and 1980s. A recent study estimated that global budgets in certain countries lowered inflation-adjusted spending on health care services by 9–17% (154). The most successful countries with global budgets for physician services are Canada (157) and Germany (19), and the most successful countries in using global budget for hospitals are those countries that implemented a case-mix-adjusted budget. Expenditure caps were also reported effective in controlling cost. For example, per admission cost in Rhode Island was 16.8% above the national average in 1980, but after the implementation of an expenditure cap, the cost per admission was 7.6% below the national average in 1992. Similarly, while expenditure per patient per day exceeded the national average by 6.2% in 1980, by 1992 it was 2.3% below the national average for community hospitals (159). Because of its success in Canada, Germany, and other OECD countries in controlling costs, the global budget is gaining increasing attention in other parts of the world (98). It was adopted as a strategy of cost control in the health care reform plan of the Clinton administration in the United States (160). Though it is effective in controlling the cost of care, efficiency cannot be achieved unless the quality of and the access to necessary care can be ensured; and allocative efficiency cannot be achieved unless the change in provider behaviour leads to the provision of more cost-effective interventions and fewer cost-ineffective interventions.

In a review of implementation of global budgets in various countries, Wolfe et al. find that although some reports claim that the global budget has an effect on cost containment, the literature is largely descriptive. It presents little rigorous empirical assessment of the effects of the global budgeting schemes in comparison to other alternatives, because global budgets are typically employed as elements of a country's overall approach to financing health benefits and controlling expenditures, and are not structured as experiments that would permit evaluation. The authors found no analytical literature that attempted to quantify the effects of global budgeting on cost, quality and health outcome (154).

The supposed virtues of the global budget have been challenged by some findings. As German ambulatory care physicians divide a constant budget over increasing services, fees decline;

physicians then provide more visits to maintain income, resulting in a rate of visits per capita more than twice that in the United States. Also in Germany, some office-based specialists are constrained in their introduction of new technologies because the nationally determined relative fee schedule cannot adjust quickly to changing expenses. Canadian hospitals have been known to close between Christmas and New Year for all but essential services, in order to remain within their budgets. Dutch hospitals develop internal budgets to match costs to resource allocations; and if budgeted supplies for a certain procedure or operation are exhausted by mid-year, hospital managements curtail that service to contain the total budget. In all these examples, the incentives transmitted to providers encourage behaviour that is most likely to move care provision away from, rather than towards, the mix most valued by society (161).

Performance-related pay

Performance-related pay (PRP) means that payment is directly linked to the performance of health care providers. PRP can be used to pay individuals or a group of people by an organization, or to pay an organization by some other agency. Performance means how well a defined task is implemented against a set target, according to the objectives of the payer.

Although PRP has existed since the 18th century (see Box 1), it was formally introduced into the health sector at the end of the 1980s, and its application is spreading rapidly. PRP is used to pay health care organizations, managers, nurses, dentists, hospital-based doctors, office-based doctors, and laboratory technicians. PRP for managers was implemented by the United Kingdom at the end of the 1980s. A total of 1400 general managers and board level mangers in regions and districts within the National Health Service (NHS) system were covered by PRP schemes in 1988, and more than 7000 additional middle-level managers of the NHS were entitled to be covered later by PRP schemes (162). PRP for nurses and nursing executives was widely reported in North America and the United Kingdom (163-165). The introduction of PRP to dental practice in the United Kingdom was reported by Scola (166). Application to hospital-based doctors was reported in the United States (167,168) and in the United Kingdom (169-175). There are some cases in which office-based doctors are rewarded based on their performance, either by a third party (176,177) or by their employer (178).The utilization of PRP for a hospital clinical laboratory was reported in the United States by Winkelman et al (179).

Box 1. Payment by performance: England in the 18th century

For reasons not difficult to understand, doctors have never favoured payment by performance, where the magnitude of their fee, if any, is proportional to their success or otherwise in curing the patient. But payment by performance contracts were not unknown in England between the 16th and 18th centuries, sometimes being imposed on unwilling physicians, surgeons, or apothecaries by cost-conscious parish churchwardens responsible for disbursing poor rate monies for the treatment of the sick, among the "deserving poor". Here, for example, is a payment by performance agreement, dated 1723, found among the documents in the parish chest at Cuckfield, West Sussex.

MEMORANDUM

An agreement made between We whose names are underwritten all inhabitants of the parish of Cockfield and George Mace of Cockfield Apothecary this 27th day of December 1723.

First We the Inhabitants have agreed to pay George Mace the sum of Four Pounds and Four Shillings in case he makes a perfect Cure of Thomas Bashford's Legg and Foot before Easter next.

In Case the said George Mace does not make a Cure of the said Thomas Bashford's Legg and Foot before Easter next, then we agree to pay him Four Pounds and Four Shillings within a year after he shall have made a perfect Cure of the said Bashford's Legg and Foot.

But in case the said George Mace shall make a Perfect Cure of the said Bashford's Legg and Foot before Easter next and shall have received the Four Pounds and Four Shillings for so doing and the said Bashford's Legg and Foot shall happen to grow bad again within a year, it is agreed that the said Geoge Mace shall repay the said Four Pounds and Four Shillings into some of the parishioners' hands for the parish use.

Witness our Hand
Robt Norden Charles Savage Berd Heassman
Mich Field — Churchwardens
Walter Gatland William Anscomb — Overseers
Geoge Mace
(Adapted from: BMJ, 1996 Volume 312, page 1453)

While PRP is being widely introduced in health care, there is no definite evidence of its effects. Opposition and support are mostly derived from theoretical arguments rather than empirical evidence. Several advantages are identified by proponents: PRP can be used to reward a good contribution (combination of effort and capability); it forces payers to evaluate their payees; and most importantly, it motivates people to perform better (180).

It is interesting to note that while health care administrators and management are often very keen to introduce the PRP schemes, academic scholars show a great deal of opposition. They claim that health care lacks the basic requirements to undertake PRP (171). First, employees must believe that improved performance will be rewarded. That means that the health care

organization must have the financial capacity to pay for the employees' better performance and the pay must be equitable. But because of the limitation in resources and the pressure for cost control, health care employers have to choose whether to reward a handful of high performers or spread the performance payment more thinly across the whole workforce. The first choice will make the majority of the employees doubt the link between performance and reward, while the second choice will lead to an almost equal distribution of the bonus budget and provide little incentive to better performance. Second, performance must be measurable and clearly attributable to individuals. Since in health care, cooperation between medical staff is needed for better quality, better performance is usually the outcome of joint efforts; and the performance of the individual medical staff is difficult to measure. There are a number of jobs where the measurable is not meaningful and the meaningful is not measurable. For example, the number of night shifts taken over by a doctor may represent his or her work effort, but it may not necessarily mean better performance or better quality; the health outcome of the patient is the most sound evidence of better performance, but in most cases it is difficult to measure. Third, rewards must be large enough to be valued by the medical staff. In a high-income country where doctors already receive high pay and the labour cost of medical care accounts for a dominant percentage of medical expenditure, the significant additional pay will lead to more serious cost inflation. For example, in the United Kingdom, pay accounts for about 75% of the cost of NHS (170). In a low- or middle-income country where doctors are paid almost equally to comparable disciplines, significant additional pay will lead to higher earnings of medical doctors and performance may not improve significantly. Griffin also argues that payment is just one of the factors that motivate the medical profession. Participation, job enrichment, recognition, decision in allocating resources, working environment, and so on can be equally important in motivating people (171). Martin's discussion (181) provides little comfort for managers and others who wish to introduce a form of individual PRP into nursing, simply because the prerequisites for a successful PRP scheme do not exist. Reviewing the literature, Lemieux-Charles states that the nature of the performance in medical care to be evaluated is ambiguous, and that there are still challenges of developing standards, guidelines and policies as well as defining quality in relation to performance (182).

Indeed, the measurement of provider performance is a major challenge to the implementation and effectiveness of PRP, not only because of the technical difficulty of measurement, but because the choice of how to measure provider performance should have a direct impact on the performance of the health system. A good performance may mean different things to different people. From the point of view of a hospital manager, good performance may mean financial viability and hospital productivity. From the point of view of health policy-makers, hospital performance means high quality, good health outcome and efficient use of limited health resources. And from the point of view of third party purchasers, good performance means reduction in utilization, cost saving, and satisfaction of beneficiaries.

Hospitals with different objective functions may design their PRP schemes differently. For example, in China the financially autonomous public hospitals which are paid on an FFS basis are increasingly using quantity of services provided or service revenue generated as a measure of the performance of hospital-based doctors. As a result, doctors have economic incentives to prescribe more and costlier drugs and recommend more and costlier medical procedures. Around 20% of the expenditure for services and drugs provided is medically unnecessary (183). The use of incentives is widespread in health care in the United States. The 1998 Physicians Benchmarking Survey data show that 64% of organizations use a salary plus bonus plan to compensate employed physicians and partners. The bases of bonus include gross professional

charges, net collection of revenue, ancillary service revenue, the number of patient encounters, and patient satisfaction (184).

Health economists who have studied physician practice patterns have generally concluded that physicians can and will alter both the volume of services produced and their use of ancillary services in response to financial rewards (14). Whether PRP can yield a socially desirable result depends on whether the performance measures used by health care providers or third party payers are in line with the desirable performance of the health system. While the socially desirable measures of provider performance should include quality of care, health outcome, efficiency, and appropriateness of medical practice, providers are very keen to include quantity of service provided and service revenue collected. Under different third-party payment systems, hospitals may design a PRP scheme that is in line with their financial objectives, but that results either in overprovision of unnecessary services (in the case of FFS) or in underprovision of medically necessary services (under capitation payment), both of which can damage the performance of the health system. It is inappropriate to say generally whether PRP is bad or good, because the goodness or badness of this payment system depends on how it is designed. Although it is possible to design and use PRP based on quality and health outcome to pay providers, as suggested by Kane & Chapin, evidence of the effect of particular PRP schemes on cost, quality and health outcome is hardly available (185,186).

Rate setting

Rate setting means the regulation of the level of fee for health services and medical products. Rate setting has been used by many countries, with a major objective to control costs. Controlling the prices of health interventions is a straightforward means to reduce spending relative to the more complex task of reducing the quantity of services or developing policies that would directly regulate quantity of care and future technological change. Prices are easy to measure and monitor. Thus, policies that affect the prices of interventions are perceived by many as the most manageable approach to achieving control over health spending (187).

Rate setting or price control in the United States generally showed an effect on the containment of costs, although the control of price led to an increase in the quantity of services provided (188). The prices of health services were frozen during the Economic Stability Program in the early 1970s, and the Medicare programme imposed a freeze on physician fees during the mid-1980s. Under the Medicare programme, hospitals have been paid a prospectively determined amount per admission based on diagnosis since 1983; physicians are paid under a fee schedule based on a relative value system; and the levels of fees are regulated by the government (187). Mandatory rate-setting gained popularity in the United States in the 1980s in the majority of states. All-payer rate setting for hospital services has been used in several states. All studies had showed that growth in expense per day and per admission was substantially less during the late 1970s in states with mandatory rate setting than in the remaining states. All studies using regression analysis to control for many other factors showed that rate-setting programmes reduced the growth of hospital cost inflation by about 3–4 percentage points a year relative to other states (189-191).

In Japan, health care service provision is integrated with drug dispensing. Health care providers are paid based on FFS and are allowed to sell drugs at a regulated mark-up. These mechanisms provide incentives for the providers to provide more services and prescribe (sell) more and

costlier drugs. Nonetheless, Japan has been able to contain its health care expenditure at a reasonable level. From 1980 to 1990, health care expenditures as a percentage of GDP increased only slightly from 6.6% to 6.8%. Although the cost containment was partly a result of the economic growth in the 1980s, the main reason for the successful cost containment lay in the government's tight control of the fee schedule. A study showed that costs had been contained mainly by reducing the level of fees for procedures, particularly laboratory tests that showed an inappropriate increase in volume, and by reducing the price of drugs that were found to have been purchased by providers at large discounts (192).

In Japan, payment to all providers is made on a uniform fee schedule regardless of geographical location and whether the providers are hospitals or clinics (193,194). The fee is inclusive and covers all supplies and materials, capital depreciation, and personnel costs. Payment for any given service is fixed, with no consideration of the actual cost of provision. Hospitals and clinics cannot bill patients extra, except for room and board and very limited specialized services. Since there is very little service being paid from outside the social insurance system, this payment system has been of crucial importance in keeping expenditures down and determining the use of health resources. Negotiation over the fee schedule occurs in the central Social Medical Care Council of the Ministry of Health and Welfare every other year, with negotiators representing providers and the public interest. Increases in physician fees were tied to increase in per capita GNP, defining a total budget cap in the negotiators' minds. The prices for pharmaceuticals are based on market price surveys (e.g., the price of drugs sold by pharmaceutical companies to drug sellers at heavily discounted rates will be reduced). Experience in Japan showed that the control of prices of services had an effect on cost control, although there was an observed change in provider behaviour — an increase in provision of services and products that were more profitable (195).

Not all countries that have controlled prices have, however, been successful. The continuous rise in the cost of drugs has caused increasing concern to the governments of European Union countries since the 1970s. Expenditure on drugs has risen faster than GDP in all member states of the European Union in the past 15 years and currently accounts for 10 to 20% of the total cost of health care (196). Regulation of drug prices has been used as a major measure to control drug costs in most of these European Union countries. There are three major strategies to control drug prices. The first is direct regulation of drug prices based on costs and effectiveness, as in Australia, Finland and Greece. The second is the use of reference prices, by which a drug is given a reference price above which the insured have to pay out of pocket. Reference drug pricing has been implemented in France, Germany, the Netherlands and Sweden. The third strategy is the regulation of the profit level of the pharmaceutical company. Prices are set by the pharmaceutical industry and are indirectly controlled through the regulation of drug profit. This is implemented exclusively in the United Kingdom.

Direct regulation of drug prices has not proved to be effective in controlling drug cost, although strict control systems have been the major cost-containment measure in several countries in the past 20 years. The direct regulation of drug prices in France has not prevented a significant increase in the volume of drug consumption (196). Two studies confirmed that direct price controls failed to contain drugs expenditure in Spain and Sweden. In Sweden, although the relative price of drugs decreased by 35% from 1974 to 1993, real drug expenditure increased by 95% and the number of prescribed drugs increased by 22% (197). In Spain, although drug prices decreased by 39% from 1980 to 1996, real expenditure for drugs increased by 264% and the

number of prescribed items increased by 10% (198). The main reasons for the increase in drug expenditure with a decreased drug price are the introduction of new drugs that are not necessarily innovative but can be charged at a higher price, and the increased quantity of drug use. In the United Kingdom, the profits of the pharmaceutical industries have been regulated since 1969, but recently this approach was criticized as having no robust reason for the regulation, either theoretical or empirical, no due process, and no impact analysis (199).

The use of reference prices for drugs has two possible effects: control of drug prices and control of drug expenditure. The effect of reference prices on the control of drug expenditure is still controversial. While there is a finding that the reference price system for drugs saves a considerable amount in Germany (200), another report from Germany states that the reference price scheme is an effective tool for price control, but its effect on cost containment is illusive (201). The experience of many countries showed that effective control of drug expenditure requires not only regulation of drug prices but control of quantity as well. The failure to control the quantity of drug use may be a major reason for the failure to control drug expenditure.

In the absence of any changes in the quantity and mix of services, reducing the prices of services will lead to lower total expenditures. However, studies of the effects of fee freezes or price controls suggest that changes in the quantity and mix of services typically do occur. More services are provided to offset the reduction in providers' revenue; providers substitute other services for those whose prices are controlled. Experience in various countries shows that the effectiveness of price control on cost containment depends on the extent to which the provider can increase the quantity of services and introduce new products or health interventions that are more profitable. The major shortcoming of price control is that it controls only one of the two parameters (price and quantity) of health care cost. Another factor to affect the effectiveness of price control is whether the regulated prices are enforced on all payers or only some. Price controls implemented for a specific population group may result in higher prices charged or an increase in services provided to other population groups. When prices are controlled for only some groups, those groups may have reduced access to health care or receive lower quality service than others (187). Experience in Japan and in some states of the United States showed that price control for all payers or under a single-payer system is more effective for cost control. Unfortunately, there are no formal studies on the effect of price control on the quality of care and health outcome (202).

So far as the level of fees is concerned, it is argued that if the fee is higher than cost, hospitals can compete on non-price dimensions such as high-tech equipment and amenities or advertising to attract patients and doctors. If the fee is lower than cost, then quality can be reduced to keep cost within income. If fees are regulated based on average costs of hospitals in an area, high-cost hospitals will reduce quality and low-cost hospitals will increase amenities; both behaviours will not necessarily increase efficiency. Again there is no evidence regarding how the level of fees affects quality and health outcome (203).

Fee structuring

Fee structure means the relative levels of fees among various services and products, particularly the rate of cost recovery of one health intervention in relation to others. Fee

structuring as a policy tool is the manipulation of fee structure so as to influence the behaviour of providers, and hence the quantity and type of health interventions.

It is generally agreed that the fees for health interventions should be regulated at a level that allows providers to recover their costs of provision through charges or subsidies or both. If there is no subsidy for providers or the subsidy is used only for expanding capacity, the prices of health interventions should be regulated at their average cost levels. However, given the complexity of determining the true resource costs for services, setting the right price for each unit of service is difficult. The common practice under a regulated price system is to assume that each provider serves a broad mix of patients and provides a broad mix of services (under FFS); and that services that are slightly underpaid and overpaid should cancel each other out, and providers will not be excessively penalized or rewarded (142).

This pricing strategy causes variation of fee/cost ratios among the health interventions offered by a provider. If the provider is financially independent, it is highly likely that its provision will be more concentrated on the interventions with higher fee/cost ratios than others. Allocative inefficiency is likely to occur if less cost-effective interventions have higher fee/cost ratios than more cost-effective interventions because of the economic benefit to providers, and because users are in a very weak position to make cost-effective choices because of asymmetry of information.

The distortions of regulated prices can be divided into the following types, and all of them may lead to allocative inefficiency. First, the level of fee may be relatively lower for primary care than for specialist care. The Resources-Based Relative Value Scale (RBRVS) provides a set of relative values for paying for physician services by taking into consideration total work input by the physician, practice cost, and the cost of speciality training (204). Before this was applied for the payment of physician services, physicians had been paid based on Customary, Prevailing and Reasonable (CPR) fees, according to which the insurers had tended to value services provided in institutional settings and technology-related services, such as laboratory and radiological services, more generously than ambulatory services. For example, the 1964 California Relative Value System, employed by more than one-half of Medicare carriers, assigns a relative value unit of 1 to a routine office or hospital follow-up visit, 80 units to a reduction of a fracture, and 1.2 units for a complete blood count. An ordinary office follow-up visit requires 13 minutes of a physician's time, whereas treatment of a fracture requires 120 physician minutes, and a complete blood count may require less than 1 minute of a physician's time (205).

Second, the level of fee may be lower for rural physicians than for urban physicians. In many countries, fees in urban physician-rich areas historically tended to exceed those in rural physician-poor areas. Under FFS, to encourage physicians to practise in underserved areas, fee schedules could potentially be set relatively lower in physician-rich areas than in physician-shortage areas (205). Third, in the absence of cost-effectiveness evidence, high-tech services are usually priced at a higher fee/cost ratio than ordinary services. This is widely observed in low- and middle-income countries, such as China (206) and Lebanon (207). This type of distortion has led to overprovision of high-tech services and underprovision of ordinary and probably cost-effective services. Fourth, new and patented drugs are priced at higher profit margins, as in the case of Japan (195). This was reported as one of the reasons why some new drugs with dubious evidence of cost-effectiveness entered the market and gained an increasing share of sales (192).

When distorted prices of health interventions result in allocative inefficiency, the correction of such a distortion is expected to improve allocative efficiency. For example, in the United States, the federal government is already using economic incentives to affect the number and mix of services. The RBRVS used by Medicare was in part adopted to encourage physicians to choose or stay in evaluation-oriented areas rather than procedure-oriented specialities, and to practise in less urbanized areas, so that more primary services and fewer specialist services could be provided (208). In the United Kingdom, general practitioners are generously paid for preventive care based on a FFS incentive to provide cost-effective preventive services. In most European Union countries, pharmacists are paid a flat rate for dispensing each prescription regardless of the cost of drugs in order to encourage them to recommend more cost-effective generic drug substitutes. All these policies are expected to have a positive effect on allocative efficiency.

The most successful story is reported from Japan, where the government not only has exerted a tight control on the price of medical services, but also has been promoting the provision of certain services through manipulation of their prices (193,194). The fee schedule in Japan is based on a point system, in which each service is given a certain number of points. In contrast to many other countries, such as Canada, Germany and the United States, rather than changing the conversion factor (which has been stable over the past 20 years), Japanese fee negotiators have been changing the points (or relative values) of the services. In this way, Japan is able to promote the provision of certain services through higher fees and reduce the amount of some services through lower fees. Historically, the increase in the fee schedule has tended to favour the services performed by clinic-based physicians over hospital services, and to encourage the provision of selected services, such as primary care.

While it is quite likely that appropriately structuring fees can lead to improvement of allocative efficiency, whether this is the case depends on whether the increased interventions are relatively cost-effective and the decreased interventions are relatively cost-ineffective. Lack of evidence on these constitutes a barrier to conclusions.

Separating dispensing from prescribing

Because expenditure for drugs represents a significant share of total health care expenditure, various methods for controlling inappropriate drug consumption have been tried in many countries. These include controlling drug prices in European Union countries and Japan, cost sharing for drugs, and the use of medical practice guidelines and drug formularies to influence doctors' prescriptions in many countries. There is little evidence, however, on whether these measures have reduced inappropriate drug utilization and costs. In recent years, debate has intensified on the possible motivation for doctors to overprescribe medication when providers are allowed to sell as well as to prescribe drugs. Theoretically, providers' incentive to prescribe is of key importance. If there is a link between providers' revenue or income and the quantity or monetary value of drug prescriptions, providers will be likely to prescribe more and costlier drugs. In countries where earning a mark-up for selling drugs is a way to compensate health care providers, the use of drugs and the percentage of drug expenditure in total health expenditure should be high. This is exactly the case in China and Japan, where the integration of prescription and dispensing has led to a larger share of drug expenditure than in countries where providers are not allowed to sell medications.

In Japan, although drug prices have been tightly controlled by the government and the reduction in prices allows little potential for the pharmaceutical companies to provide significant discounts, the integration of prescription and dispensing provides an incentive for doctors to prescribe. Before the 1990s, the Japanese consumed large quantities of prescription medications: Japan ranked second only to the United States in the size of its total drug market. On a per capita basis, Japan had the highest medication use of any country in the world. In 1988, drug expenditure in Japan was $332, much more than the $182 in the United States (209). Drug revenues of the inpatient services accounted for 29.2% of total revenue; and drug revenue accounted for 44% of the total outpatient revenue (192). Much of this medication use may be the result of the close link between prescribing and dispensing by physicians. Although there is no scientific documentation to support such a claim, Japanese medication use may be higher than necessary.

In the 1990s, the Japanese government introduced controls such that the profit margin for less effective drugs was lower than for more effective drugs. This policy change makes it less profitable for doctors to dispense drugs. The national health accounts data for 2000 show that Japan is only eighth in the world in terms of total pharmaceutical expenditure (in international dollars) per capita.

The dispensing of drugs in China is the same as in Japan. The profits on drug sales by providers are an important source of their income. There is, on average, a 20% mark-up on medication sales. It was estimated that the profit rate was no less than 15%. National statistics showed that in the early 1990s about 60% of hospital revenue and more than 80% of the revenue of practice doctors was generated through selling drugs (210); and the expenditure for drugs accounted for more than 40% of total health care expenditure, the highest share in the world (211,212). It was reported that about 30–40% of the drug consumption arose from overutilization or inappropriate use (183,213). Starting from 2000, China has made efforts to remove the incentives for overprescription. Two measures have been taken: one is to reduce the prices of drugs through bidding for acceptance by the social health insurance schemes at city level; and another is to request public hospitals to set up independent cost and revenue accounts for drugs, with the profit from selling drugs to be managed by government. These reforms are still in process, and the effects are yet to be evaluated.

In African countries, charges for drugs (at a mark-up of more than 100%) in publicly funded community health centres resulted in many prescriptions containing injections (56%) or three or more drugs (89%). These prescriptions yield 120–200% profit to the health centres, at a social cost of irrational drug utilization (214). In Nepal, costsharing for drugs led to more poly-pharmacy and excessive drugs use (215).

In the comparative studies in Zimbabwe (216-218), which involved 28 private sector dispensing doctors, 25 non-dispensing doctors and a total number of 688 patient records, the authors found that dispensing doctors prescribed significantly more drugs, more injections and more antibiotics per patient than non-dispensing doctors. Compared with non-dispensing doctors, dispensing doctors used subcurative doses of cotrimoxazole more often (26.4% of the encounters compared with 11.7%) and correct doses less frequently (58.0% compared with 72.6%). On average, dispensing doctors prescribed significantly more drugs per patient than non-dispensing doctors (2.3 versus 1.7), injected more patients (28.4% versus 9.5%), and prescribed more antibiotics (0.72 versus 0.54) and mixtures (0.43 versus 0.25) per encounter. Using the same dataset, the authors analysed prescriptions by dispensing and non-dispensing

doctors for patients diagnosed with upper respiratory tract infection. It was found that dispensing doctors were associated with a greater number of drugs per encounter, a greater number of injections, and more use of analgesic drugs, cough and cold preparations and psycholeptics. Multivariate analyses controlling for sex, race, place of education, location of practice and patients seen per day showed that dispensing by doctors was associated with less clinically and economically appropriate prescribing. These findings suggest that the quality of health care — as related to drug use, patient safety and treatment cost — is lower with dispensing doctors than with non-dispensing doctors.

The belief that the integrated system of prescribing and selling is a major reason for increased drug use has led to the intention of reforming the system. In Japan, independent pharmacies have been set up and patients are encouraged to obtain drugs from these pharmacies. In Taiwan, the mark-up rate for pharmaceuticals has been reduced in order to decrease the economic incentive for the provider to sell more and costlier drugs. In China, policy-makers are discussing ways to reform the system and remove the incentive for inappropriate prescribing. In the Republic of Korea, a proposal to separate dispensing from prescribing has been worked out. While there is evidence that dispensing providers often prescribe more and costlier drugs than non-dispensing providers, there is little study of the effects of the reform of separating dispensing from prescribing on the costs of drugs, quality, health outcome, and the mix of drug prescription and utilization.

Payer integration

Payer integration is a policy that integrates multiple payers of health care into one or several major payers as so to control the overall health expenditure of a country. The major rationale is that in a health care system where a provider's revenue comes from multiple sources, and insurers pay providers in an unorganized manner, with different methods and at different rates of payment, it is difficult to control the overall expenditure on health care. Restriction on payment from one source may lead to cost shifting to other sources; and savings resulting from payment reform for one insurer may lead to increase in costs for others. At the extreme, under a single-payer system with revenues of all providers from only one source, control of payment by this source implies control of the total amount of health expenditure of a nation. The cost shifting or squeezing effect can be eliminated. In contrast, the larger the number of payers, the more difficult it is to harmonize the flows of funds from different expenditure sources, and the more difficult it is to control the overall expenditure of a country.

The integrated payer system has existed for a long time, but it was not recognized as a policy tool until the increase in health care cost become a serious question in the United States. During the debate over reform of the United States health care system in the early 1990s, a single-payer system was proposed as one of the reform options for the achievement of universal coverage and cost containment; the multi-payer system in the United States was criticized; and the advantages of the integrated payer system were discussed.

Figure 1 shows a simplified view of the organization of health care in the United States. It is characterized by private financing and private provision coupled with government-financed

health programmes. The population is covered by various health insurance schemes, including insurance provided by 1500 private insurance companies, the Medicare programme organized by the federal government which provides coverage for the elderly and disabled, and the Medicaid programme organized at state level which provides coverage of the poor. Each of the insurers, public or private, collects funds and pays health care providers (hospitals and physicians). As shown in Figure 1, one insurer pays multiple providers (salary, FFS or capitation for physicians, and per diem or DRG for hospitals); and the payment to the providers usually comes from multiple sources, including various insurance programmes, cost sharing by the beneficiaries, and full payment by patients who are not insured. This structure forms a multi-payer system in which there are multiple fee schedules, and the flows of funds are difficult to monitor and control.

This type of system suffers from two serious problems. First, because private insurance is largely employment based, if there is no government-organized insurance scheme to cover the unemployed, there will be uninsured people, the number depending the rate of unemployment. The United States has been criticized as being the only high-income country which has not mandated universal coverage for its population. It is reported that at least 37 million of the United States population lack any form of health insurance coverage (219). Second, high cost and low efficiency are inevitable. The multiple commercial insurers compete with each other using various marketing techniques, conduct underwriting on a case-to-case basis to select individual or non-group enrolees, and undertake utilization review and reimburse providers with various arrangements. All these have to incur costs that would otherwise be unnecessary. While the administrative cost for the single-payer Medicare programme was only 2% of Medicare expenditure, administration accounted for 20–24% of the total expenditure of commercial insurance. While the United States health care system spent 20% of its total health care expenditure on administration, the Canadian single-payer system only spent 10% (220). The unmanageable multi-payer system is believed to be one of the reasons why the United States spent approximately twice as large a share of GDP as that in Australia, Canada, Japan, Germany and the United Kingdom (221).

To get a sense of what an integrated payer system looks like, Figures 2 to 5 show, respectively, a simplified outline of the health care systems in the United Kingdom, Canada, Germany, and Japan. In the United Kingdom (Figure 2), the health care system is characterized by public financing and public provision; the central government serves as both an insurer and a payer; the NHS is financed through general taxation and provides universal health insurance coverage to its population; public hospitals are reimbursed based on global budget, and physicians are paid a capitated fee. Although there are payments outside the NHS system (e.g., the use of private providers and patient co-payment for drugs), the vast majority of payments occur within the NHS, so the national government is in a dominant position to control the total amount and allocation of health care expenditure (222).

In Canada (Figure 3), the health care system is characterized by public financing and private provision. Its health care financing system is the same as in the United Kingdom, but is or ganized at provincial level, with financial sources from both local and federal governments. Its health care provision system is similar to that in the United States (where private hospitals and physicians dominate the provision system, and a physician usually works in both hospitals and his or her office) and different from that in the United Kingdom (where public providers dominate, and there is a clear distinction between office-based doctors (general practitioners) and hospital-based doctors (consultants). The financing system provides universal coverage;

hospitals are reimbursed based on global budgets, and physicians are paid on a FFS basis under a global budget for all physician services. As in the United Kingdom, although there are payments outside the socialized health care system, the majority of the payments happen within the system, and the provincial government is in a dominant position to control the total amount and allocation of health care expenditure (223).

In Germany (Figure 4), the health care system shares the main characteristics of the United States system — private financing and private provision. The health insurance schemes in Germany are organized by various insurers (sickness funds) with financial contributions from employees, employers and government subsidies. Germany has a provision system dominated by private providers. In contrast to the United States, universal coverage is realized through tight regulation by government, which requires mandatory insurance coverage and financial contributions from various sources based on ability to pay. Another major difference is that the sickness funds pay the state medical associations yearly, which in turn compensate physicians based on a relative FFS schedule which is negotiated annually between the sickness funds and the national medical association. A global budget for physician services is regulated by each state; payment to physicians by the state medical association is based on the negotiated fee schedule and a global budget. Hospitals were paid on a per diem basis before 1992. The per diem rate was negotiated between individual sickness funds and individual hospitals. The per diem fee covered only the operating cost of the hospital, and capital expenditure came from the government global budget. Since 1992, hospitals have been paid based on a combination of case payment, procedure fees, and per diem; the rates of payment are regulated by federal and state governments; and government pays the capital cost. Because of the integration of payment to physicians by medical associations (rather than by individual sickness funds), and because the insurance payments account for a dominant part of the total expenditure for physician services, the expenditure for physician services is under the control of the government. For hospital expenditure before 1992, capital investment — a major part of hospital cost — was under the control of government which had financial responsibility to provide investment funds within a global capital budget. Since 1992, the fee schedule for hospital services has covered all payers, and payment from outside the health insurance system is limited. Total hospital cost can be controlled through the control of medical fees (19,224).

In Japan (Figure 5), the health system is quite similar to that of Germany and the United States — private financing and private provision. Various health insurance schemes are organized by private entities and government, with financial contributions from employees, employers and government subsidies; provision is dominated by private providers. As in Germany, and differing from the United States, universal coverage is realized through mandatory insurance coverage and financial contributions from various sources based on ability to pay. Unique to Japan is that both hospitals and physicians are paid based on FFS with an all-payer fee schedule which is negotiated at central level approximately every other year. Also unique to Japan is that two organizations (the National Health Insurance Organization and the Social Insurance Medical Care Fee Payment Fund) have been created to serve as intermediaries between the providers and insurers and to be responsible for utilization review and reimbursement. Insurers do not reimburse providers directly. Because providers are prohibited from charging patients directly and billing outside the insurance system is not allowed, the government is in a good position to control total health expenditure through its all-payer fee schedule, manipulation of fee structure, and the integrated payment system (194).

All these countries, except the United States, share three characteristics. First, payments for providers are integrated into either a single payer or only a few major payers. Second, there exists a fee schedule (negotiated or regulated) which covers all payers. Third, there is an explicit (e.g., in the United Kingdom) or implicit (e.g., in Japan, where the increase in fees is kept at a rate no more than the rate of economic growth) global budget. Apparently, the first characteristic is a prerequisite for the last two. These characteristics allow for macro-control of health care expenditure. This is different from the situation in the United States, where cost control measures are adopted at micro-level (e.g., case-by-case utilization review by individual insurers) but the overall cost is not controlled.

While payer integration is believed by researchers to be a powerful tool to control health care costs, and advocates in the United States have promoted the single payer system as an element of US health care reform (221), studies of the effect of payer integration on quality of care and health outcomes are rare, and evidence of its effect on allocative efficiency is hardly available. At macro-level, the similarity in health status and variation between countries with an integrated payer system and those with a multiple payer system suggest that countries with many payers tend to spend much more for the same level of health, and it is highly likely that health resources are not allocated in an efficient manner. Comparison between Japan and the United States found that US providers apparently provided much more elective tests and surgeries than their Japanese counterparts (193). Comparison of health care utilization among eight OECD countries showed that the number of outpatient visits per person per year and the number of hospital days per person per year in the United States are the lowest, while the use of high-tech services (e.g., MRI and CT) is the highest (9). Comparison of medical technology in Canadian, German and US hospitals found that US hospitals provide significantly more sophisticated equipment per million persons than is available in either Canada or Germany (225). The reasons for these findings must be multiple, however, and it is hard to say there is causality between the degree of payer integration and the profile of care provision.

Figure 1 Simplified structure of the health care system in the United States

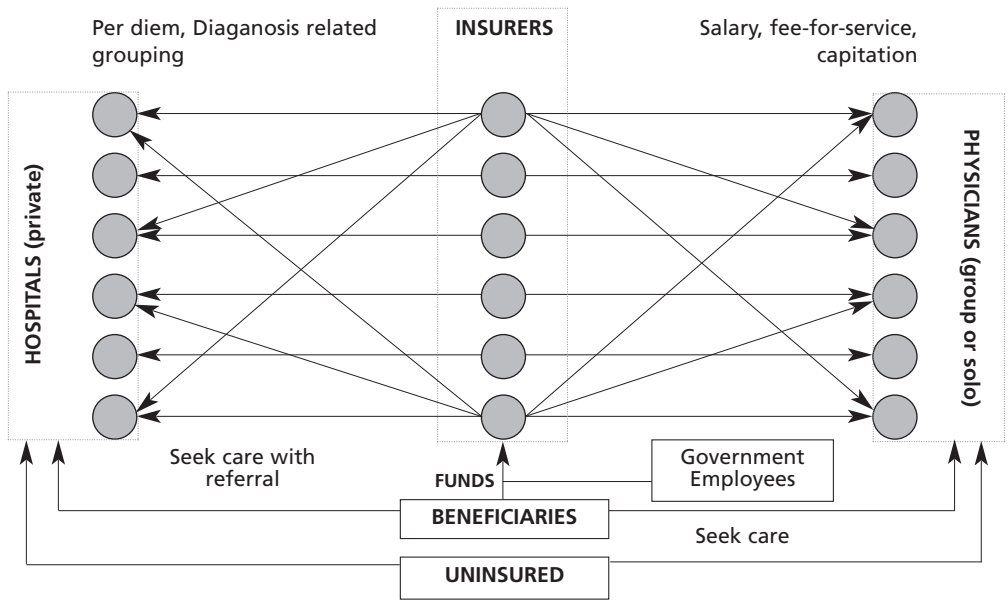


Figure 2 Simplified structure of the health care system in the United Kingdom

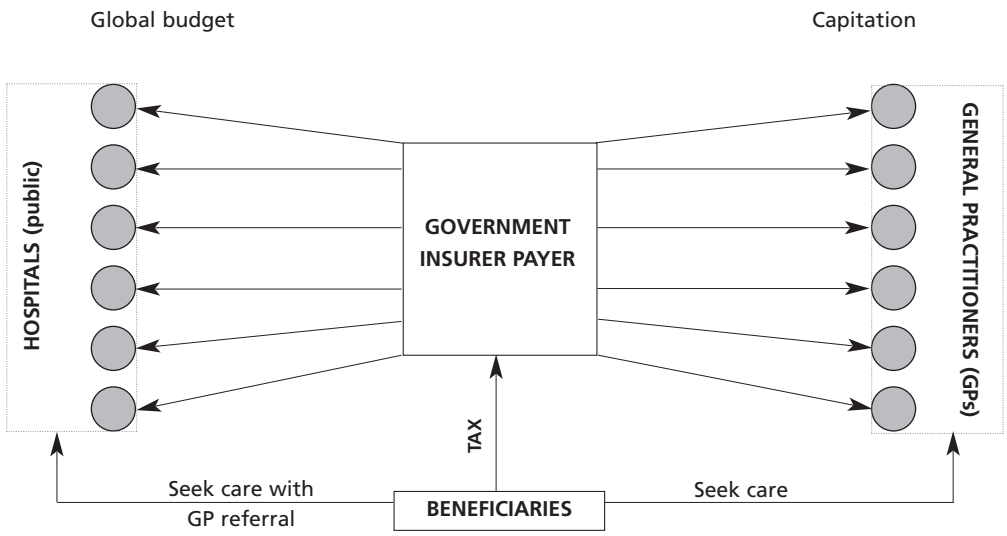


Figure 3 Simplified structure of the health care system in the Canada

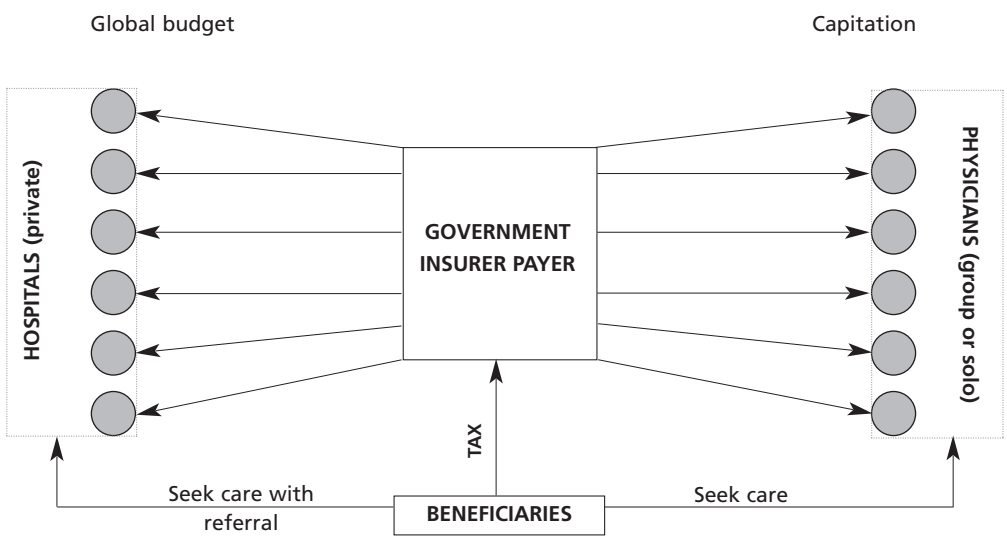


Figure 4 Simplified structure of the health care system in Germany

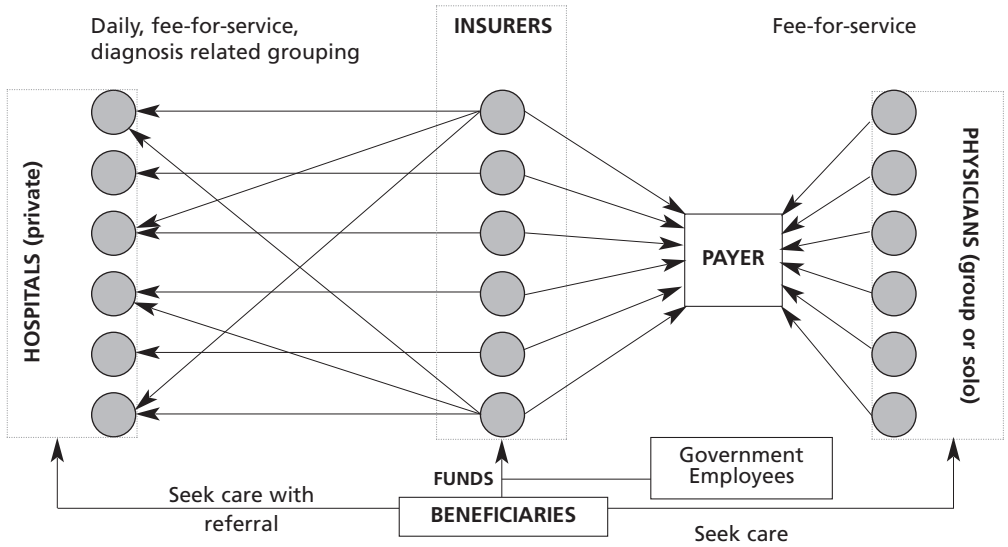
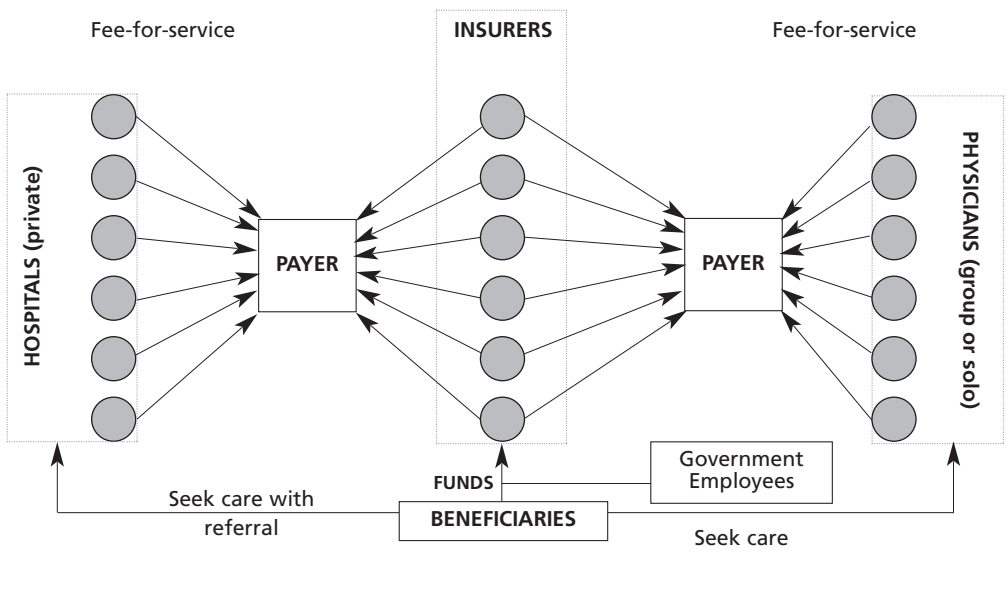


Figure 5 Simplified structure of the health care system in Japan



Chapter 4. Economic Incentives for Users

Economic incentives for users means no more than increasing their financial responsibilities for the use of health interventions. The argument for this is that if all health interventions are provided free of charge at the point of service, consumers are motivated to use more and costlier interventions even when fewer and less costly alternatives with similar effects are available. This can lead to either cost escalation of health care in a health insurance-dominated health market or shortage of supply in a public provision-dominated market. Policy tools include user fees and cost sharing.

User fees

User fees are charges levied on users of public sector health services with the aim of recovering some or all of the costs (226). Traditionally, many low- and middle-income countries provided free care to their populations, and the publiclyowned health service facilities were directly financed by governments. User fees as a component of the reforms are related to structural adjustment, downsizing the role of government in public financing and promoting market mechanisms. High population growth and budget constraints limit the ability of some governments to improve health services available to the population and even to maintain their current level. In addition, it becomes difficult to raise revenues to finance recurrent costs in public facilities. There seems to be a need to introduce user fees at an institutional level to complement shortfalls in public funding.

User fees were first proposed for low- and middle-income countries in 1985 (227); the rationales are as follows (228). First, insufficient government funding to the health sector led to deteriorating quality and accessibility of health services needed by the population; imposing user fees could generate additional revenue for the public facilities to improve quality of and people's access to basic care. Second, the publicly funded system was misused, resulting in overutilization of unnecessary care by people with more power and underutilization of necessary and cost-effective services by others; imposing user fees could deter the unnecessary use by the overserved, and encourage the use of cost-effective services by the underserved through a fee exemption system. Third, internal inefficiency led to wasteful public programmes of poor quality; user fees could make the providers more accountable for their performance and thus promote productivity of publicly owned health facilities.

Based on studies in the Philippines, the World Bank stated that demand for health care is relatively price/income inelastic, that is, people will pay for health care even though they are relatively poor (228). It is argued that charges will deter the consumption of unnecessary care and at the same time, if the quality of care provided improves, this in turn will lead to better access to care (227,229). This reform has, however, raised a great deal of concern over the possibility of decrease in utilization of necessary and cost-effective care by the people who have no ability to pay. This is both an efficiency and an equity issue.

A study by Gertler et al.demonstrated that, while demand for health care is indeed inelastic for most income groups, people in the lowest income quintile are highly responsive to price changes. While user fees may be a good way to raise needed funds, the poorest people would

be vulnerable and their access might be impeded (230). Several empirical studies have been conducted, mostly facility-based longitudinal observations without a control group. The studies in Ghana (231), Swaziland (232), Zaire (233), the United Republic of Tanzania (234) and Kenya (235) showed that the introduction of fees led to a significant decrease in utilization of facility services. In the United Republic of Tanzania, the introduction of user charges in public facilities resulted in a 50% reduction in utilization, part of which switched to private hospitals and part of which was deterred. It was not clear, however, whether the deterred utilization was unnecessary. In Kenya, the introduction of user fees in 1989 led to a 38% drop in utilization of public facilities, and later the user fees had to be cancelled because none of the poor were exempted from fees, which led to a reduction of utilization of both necessary and basic health services (235). There were additional reports that fee systems had resulted in sustained reduction in utilization and deterred those whose health care needs were greatest (236-241). These studies consistently concluded that health facilities which introduce user fees will experience a drop in overall utilization, including the use of both necessary and unnecessary care, and that the poor will be hit the hardest (242).

Nevertheless, these studies did not consider the issue of what will happen if there is an improvement in quality of care and if poor people are exempted from paying fees. In contrast to the studies described above, experiences reported from facility-based longitudinal studies in four other African countries — Benin, Sierre Leona, Guinea (243) and Cameroon (229) — showed that when fees were accompanied by a notable improvement in quality of care, overall utilization did not decrease but in fact increased. The study in Cameroon was a controlled experiment conducted in five health facilities, three of which were selected as treatment centres, and two comparable facilities not imposing user fees were selected as controls. Instead of using facility data, two rounds of household surveys were conducted to measure the percentage of ill people seeking care at the health centre before and after the implementation of the user fee policy. Results indicated that, because of the improvement in health care quality, the probability of using the health centre increased significantly for people in the treatment area as compared to those in the control areas. The access to heath care for the poor was improved as a result of introducing fees and improving the quality of care as measured by the reliability of drug supply. A study conducted in Zimbabwe found that 31% of the users were exempted from charges for both consultation fee and the costs of drugs; the poor were effectively subsidized by the rich who were able to pay, and there was no significant reduction of use for those who were exempted.

The improvement in quality and the effective implementation of an exemption system are, however, always questionable. According to an international survey (244), in 27% of the 26 countries surveyed, revenue generated was returned to the central government and not used for the improvement of service quality. In only 30% of the countries could facilities retain the revenue generated from user charges. As there were few evaluations of the impact of user fees on service quality, the conversion of revenue into quality improvement remains poorly understood. Published data from national cost recovery schemes suggested that, on average, only 5% of operating or recurrent cost was recovered, which was insufficient to address the quality problem of the health system as a whole (245).

Regarding exemptions, as pointed out by Creese, carefully discriminating fee systems are necessary to ensure that revenue is provided only by those who can afford to pay; that the resulting income improves the quality and accessibility of health care targeted at the poor; and that user charges can redistribute resources by charging the better-off to prevent unnecessary

use and subsidizing access to necessary care for the indigent (237). Empirical evidence concerning an effective implementation of an exemption system is very weak. Russell & Gilson raised the question of whether a user fee policy for promoting health service access for the poor is a wolf in sheep's clothing (244). Among the 26 low- and middle-income countries with user fees they surveyed, many lack policies that promote access for disadvantaged groups. Of those countries, 27% had no policies to exempt the poor; in contrast, health workers were exempted in 50% of the countries. Even in the countries where an exemption policy existed, implementation varied because of difficulties in identifying the real poor. The authors concluded that fees were likely to exacerbate existing inequities in health care financing. Even if the poorest are exempted, the poorer among the rest of the population will have to pay a higher proportion of their income. For example, in Zimbabwe, the same level of fees for drugs represented more than 6% of the income of the lowest income group, but only about 2% of the income of the high income group (246).

In theory, user fees may encourage more efficient use of services if the fees are graduated by level of the system (e.g., higher charges for higher levels of the system); a by-pass fee is introduced in areas where the primary care network is adequate and the referred poor patients are exempted at higher levels of the system; and user fees are associated with quality improvements which promote utilization at the primary level. There is some evidence, however, that user fees encourage overprovision of unnecessary care when the revenue is retained at the point of collection; and overall, fee systems represent weak mechanisms for improving the efficiency of utilization (247). There is no evidence that user fees can improve the mix of health interventions.

Cost sharing

Cost sharing refers to any additional direct payment made by the insured users of health care to the health care providers (248). There are three main forms of cost sharing:

- Deductible: cumulative amount that must be paid out of pocket before benefits of the insurance programme become active;
- Co-payment: flat amount that the beneficiary must pay for each instance of health care used (including services and drugs);
- Co-insurance: percentage of the total charges for care that must be paid by the beneficiary (249).

Other policies that are often associated with these include a benefit cap (the limit on the amount that can be reimbursed), an out-of-pocket cap (the limit on the amount paid by a patient or a household), and coverage exclusions (health services and drugs excluded from benefit packages). In contrast to user fees, the main objective of cost sharing is to control costs for the insurers by rationalizing demand and deterring the use of inappropriate services and drugs.

Cost sharing is widely used in US commercial health insurance. A total of 90% of privately insured individuals under indemnity plans are subject to cost sharing requirements, and over 75% of managed care enrolees face some form of cost sharing (250). Medicare beneficiaries are subject to variable amount of deductibles (about US$ 600) for each inpatient stay, and a US$ 100 deductible and 20% co-insurance rate for physician services, although around 70% of them

have Medicap insurance to cover these costs. Cost sharing for Medicaid patients, which was originally prohibited by federal regulation, is common in some states.

Nearly all countries of western Europe provide universal coverage for health care costs to their population through general tax-funded national health systems or compulsory social insurance programmes. Although it has been stated that, with the exception of France and Portugal, western European countries rely very little on cost sharing as a tool for either raising revenues or containing costs for physician and hospital services (251), a recent review of the literature (248) found that about half of western European countries use some form of cost sharing for first contact care, and about half also apply cost sharing to inpatient and specialist outpatient care. The common forms of cost sharing in European countries are co-payment and co-insurance. Only Switzerland uses deductibles. Virtually all western European countries reduce or eliminate cost sharing obligations for persons identified as belonging to a disadvantaged group. Exemption from charges is commonly based on the grounds of individual or household income and age.

Costs for drugs are rarely covered in traditional commercial health insurance in the United States, but some managed care organizations cover the cost of drugs, subject to some cost sharing arrangement. Medicaid programmes initially covered the full costs for drugs. In order to control the cost of the programmes, cost sharing for drugs for Medicaid patients has been increasingly used since the 1980s.

Cost sharing for pharmaceuticals is widespread in western Europe (252). Some countries use negative lists for drugs (such as Germany and Ireland), and some countries use positive lists (such as Belgium and Spain) to specify which drugs are in the benefit package. Pharmaceutical reference pricing has been implemented in Denmark, Germany, the Netherlands, New Zealand, Norway and Sweden. In these countries, cost sharing for drugs includes a fixed charge per prescription (co-payment), plus the difference (Germany) or a portion of the difference (New Zealand and Sweden) between retail and reference price. In France, Italy and Spain, the rates of cost sharing vary for different drugs and patients. In the United Kingdom, a flat rate per prescription (with exemptions for low income, elderly people, and patients with specific chronic diseases) is charged.

The effects of cost sharing are multiple, affecting total utilization, the mix of health interventions, the cost of health care, equity, and health outcomes.

Total utilization

All studies reviewed show that cost sharing is associated with a reduction of total utilization of services and drugs. The first notable natural experiment reported in the literature took place at Stanford University in the late 1960s. This study found that the imposition of a 25% co-insurance rate for university staff and their dependants reduced utilization by about the same percentage (253,254).

The Health Insurance Experiment (255), designed and conducted by the RAND Corporation in the 1970s and 1980s, used a true experimental design to determine the effect of patient cost sharing on the utilization and costs of medical services, and on patients' health status. This is

the most important study in the field of health services, the results of which are cited throughout the world. In general, this study found that cost sharing had an effect on the quantity of medical care demanded. The elasticity was small but definite (about -0.2), similar to the findings reported by the Stanford University study (254). Cost sharing worked almost entirely by reducing the number of medical care episodes for which treatment was sought. People with higher co-insurance rates were less likely to seek any inpatient or outpatient care.

A natural experiment occurred in 1977 in the United States. Before July 1977, the United Mine Workers health plan provided free care, and after that time large cost sharing requirements were instituted, which included a US$ 250 annual inpatient deductible and a 40% co-insurance rate on physician and most outpatient services up to a maximum of US$ 500 per family. By comparing the inpatient and outpatient utilization data before and after the reform, Scheffler found that the probability of a hospitalization dropped by 45%, and the probability of a physician visit declined by 35% (256).

In another natural experiment, Cherkin, Grothaus & Wagner examined how utilization of health care changed from the year before to the year after the introduction of a US$ 5 co-payment for ambulatory visits for HMO patients. The institution of the co-payment resulted in an 11% reduction in the number of primary care visits, a 14% decrease in physical examinations, and a 3% decline in speciality care visits (257,258).

Selby et al. studied the effect of a US$ 25–35 US dollar co-payment for use of emergency departments, based on before-and-after comparison of 30 276 HMO participants. The introduction of co-payment was associated with a decline of about 15% in the use of emergency departments (259).

A more recent study by Stuart & Zacker examined the impact of Medicaid prescription drug co-payment policies in 38 states of the United States, using cross-sectional survey data from the 1992 Medicare Current Beneficiary Survey. Elderly and disabled Medicaid recipients who resided in states with co-payment provisions had significantly lower rates of drug use than their counterparts in states without co-payments, and co-payment was associated with a 15% reduction in drug use (260).

A before-and-after comparison of a quasi-experiment with control groups showed that introduction of a US$ 1.50 co-payment per drug prescribed in a HMO was associated with a 10.7% reduction in the number of drugs prescribed. When the co-payment increased to US$ 3, there was another 10.6% reduction in the number of drugs prescribed (261).

There are some studies outside the United States which provide similar evidence. In Canada, an analysis of the impact of co-payment for physician services from 1968 to 1971 in the province of Saskatchewan showed that this caused a reduction in overall utilization of physician services (262). In Japan, co-payment in social health insurance schemes was found to have a negative effect on total service utilization (263). In Ireland, a cross-sectional household survey found that the number of visits per year for the groups with no cost sharing (5.9 for persons 35–44 years of age, and 9.4 for persons 55–64 years of age) was much greater than for the groups with cost sharing (2.7 for ages 35-44 years, and 4.0 for ages 55-64 years). Multiple regression with controls for health status and demographic variables showed that cost sharing was highly associated with lower probability of physician visits (264). In Iceland, cost sharing for specialist services was introduced in 1993; this led to a reduction in the number specialist visits of approximately 10% (265). In the United Kingdom, it was shown that a 10% increase in co-payment could lead to a 1.5–2% decrease in demand for prescriptions (266).

Mix of interventions

The reduction of total utilization does not mean an improvement in allocative efficiency, unless the reduction is mainly the result of a decrease in utilization of less cost-effective interventions. Research shows that cost sharing resulted in a reduction of both cost-effective and cost-ineffective services, and both essential and non-essential drugs; a larger reduction in utilization of cost-effective interventions than cost-ineffective interventions; and a larger reduction in patient-initiated utilization than in physician-recommended utilization.

In the Stanford University study, the analysis of utilization before and after the introduction of a 25% co-insurance in 1967 for all services covered by the Group Health Plan showed disproportional declines in utilization of different types of services. The per capita physician visits declined by 24%; outpatient ancillary services decreased by only 11%; the number of surgical hospitalizations decreased by 5%; and the total number of hospitalizations decreased by 3%.

In the RAND Health Insurance Experiment in the United States, Lohr, Brook, Kamberg et al. found that the cost sharing schemes reduced both appropriate care (highly effective) and inappropriate care (rarely effective) (267). Manning, Newhouse, Duan et al. found that cost sharing could reduce both outpatient and inpatient contacts, but there was no difference in the types and quantity of services once the patients were within the system (255).

Cost sharing of US$ 1 per visit for the first two physician visits per month imposed for Medicaid patients in California in 1972 provided another natural experiment. The imposition of a co-payment appeared to reduce utilization of certain preventive services such as immunizations, Pap smears, and obstetrical care (268). Roemer et al. stated that the imposition of copayment for Medicaid was penny wise and pound foolish. They found that cost sharing was associated with a 4% reduction in the number of visits, and also associated with an increase in hospitalization rates (269).

Limited evidence from other countries shows consistent results. In Japan, the utilization of consumer-initiated services is more sensitive to cost sharing than that of provider-initiated services. Kupor et al. concluded that inpatient utilization in Japan was much less sensitive to cost sharing than outpatient medical and dental care (263). In Canada, Bech & Horne found that the inpatient co-payment between 1968 and 1971 had no effect on either the number of admissions or the length of hospital stays because these services were physician-initiated (262). In Iceland, the introduction of cost sharing led to a 17% decline in the number of women receiving screening for cervical cancer, and only a 10% reduction in the utilization of specialist services (265). A thorough review of the effect of cost sharing for drugs concluded that rigorous experimental and quasi-experimental studies suggested that cost sharing could both reduce the use of both non-essential and essential drugs and might do more harm than good (270).

Cost containment

There seems to be no consensus on the effect of cost sharing on the containment of costs. Kutzin concludes that cost sharing does reduce utilization but does not contain costs. He argues that co-payments are very popular in the United States, and the United States has a lower rate

of physician contacts and per capita bed-days than many other countries, such as Canada, France, Germany Japan and the United Kingdom, but costs in the United States are much higher relative to GDP than in these other countries (248). Although his argument might provide little support for his conclusion, his suggestion that cost sharing alone has little effect if provider behaviour is not controlled to provide less intensive and more cost-effective care is quite reasonable.

The RAND Health Insurance Experiment suggests that cost sharing is associated with a decrease in total health expenditures. Co-insurance rates had a major effect on costs. Individuals with free care in 1984 incurred an annual medical expenditure of US$ 777 per capita, which were 23% greater than for those with a 25% co-insurance rate and 46% higher than for those with a 95% rate. However, only a small fraction of the population in each site participated in the experiment, so the reduction in use by the experiment participants would have little effect on the practice of any individual physician. As a result, the reduction in cost in the experiment does not reflect the real world situation, in which the reduction in utilization as a result of cost sharing of all patients of a physician will lead to changes in physicians' behaviour, recommending more services than the patient demanded if payment is based on FFS. The final result may be a reduction in patient-initiated demand, but not in the costs of health care (271). Because the elasticity of inducement of the providers is greater than the elasticity of demand of the consumers (272), the effect of cost sharing on cost containment can be offset by the change in providers' behaviour.

Analysing the data of the United Mine Worker health plan, Scheffler found that the decrease in the number of admissions was accompanied by an increase in expenditure per admission (256). Analysing the same data, Fahs found that the imposition of cost sharing for the mine workers changed the practice pattern of the physicians for patients outside the reform (273). This response took the form of raising their price for ambulatory care to patients who were not mine workers and increasing their inpatient lengths of stay. This response led to a 19% increase in inpatient expenditure and a 7% increase in total health expenditure for the patients not covered by the cost sharing reform. These results apparently meant that the costs saved were shifted to the other users who were not covered by the cost sharing reform, or to services which were less (or not at all) restricted by cost sharing.

The effect of cost sharing on the cost of drug use is more controversial. In Sweden the use of reference prices for drugs led to a 5% saving on government drug expenditure, but the total cost of pharmaceuticals continued to increase at the original rate because the saving for government was shifted to patients (274). A review of the international literature related to the effect of co-payment for drugs on cost control showed that the reduction in cost for drugs increased the cost for other components, such as acute psychiatric services and institutionalization in nursing homes for elderly people (270). As a result, the total cost for health care may not decrease.

All the studies of Medicare patients in the United States showed that the patients with Medigap coverage (a supplementary insurance covering costs that are not covered by the Medicare programme) increased not only the cost of the Medicare programme, but overall health care costs as well, meaning that cost sharing reduced the overall cost of Medicare patients (250). The authors argued that, if there were no cost sharing mechanism — which is very popular in the United States — it would be expected that health care costs as a percentage of GDP would be much higher.

Equity

Cost sharing imposes a big challenge for equity in financing and utilization of heath care. It seems that cost sharing for services and drugs has more impact on vulnerable populations, such as women, children and the poor. The United States and Switzerland rely mainly on private insurance and make greater use of cost sharing; they are found to have the most regressive health financing systems of the ten OECD countries (248).

The Stanford University study found that the percentage of individuals who did not see a physician during one year increased from 13% to 20% as a result of the 25% co-insurance requirement in the United States, the largest increase being for the lower-income group (253,254).

The RAND study found that co-insurance was a deterrent to low-income persons seeking health services that researchers judged to be highly effective. With co-insurance, the children of middle- and upper-income families used 85% as much highly effective care as was used by those children without cost sharing; while the children of low-income families used only 56%. Similarly, upper-income adults subject to co-insurance used 71% as much effective care as their counterparts with free care; while the corresponding figure for low-income adults was only 59% (267).

The study of the effect of HMO co-payment for ambulatory visits (257,258) found that for primary care visits, women under the age of 40 years had twice as great a reduction as men in the same age group. Women and children were most affected with respect to seeking annual physician examinations. Rates for children and women declined by 20–25% and 15% respectively, but there was no detectable change for men. Freemantle & Bloor concluded that cost sharing for drugs was also more likely to deter the use of drugs by vulnerable people (270).

Health outcomes

The evidence on the impact of cost sharing on health outcomes is rather limited. The Stanford University study did not examine the effect on health; the RAND study found that the reduction in use as a result of cost sharing had an observable negative effect on health only for the poor who had a higher risk for ill-health. Several studies reviewed by Rice & Morrison (250) compared people with insurance and those without. The people without insurance (100% co-insurance rate) had shorter hospital stays and tended to receive fewer high-cost services, and the uninsured appeared to experience more avoidable hospitalization and higher rates of adverse birth outcomes. However, most studies were not able to find a direct association between poorer outcomes and the less-intensive process of care experienced by uninsured persons (275).

Summary

This literature review finds that cost sharing can consistently lead to a reduction in overall utilization of services and drugs, but the reduction in utilization may not necessarily lead to a reduction in the overall cost of health care because of possible cost shifting from the restricted

services to those that are not, and from cost sharing people to others. In addition, the reduction in the use of essential and cost-effective services may lead to an increase in utilization of costlier health interventions. Cost sharing may inequitably affect vulnerable people such as women, children and the poor with regard to the use of preventive and cost-effective services and drugs. While the intention of cost sharing, particularly in the publicly financed health care system, is to reduce the inappropriate use of services and drugs, and thus to improve allocative efficiency, the findings from research are not optimistic. It is very common that cost sharing results in a greater decrease in essential, cost-effective health interventions.

These conclusions do not rule out the potential for using cost-sharing to control costs, and to improve equity and allocative efficiency. First, to control health care costs, the change in provider behaviour resulting from a reduction of utilization should not be ignored. Cross-country data suggest that consumer-initiated demand is not the major factor driving health care costs; instead, it appears to be the intensity of services recommended by medical doctors (276). Since intensity is largely provider-initiated, there is wide scope to control costs through supply-side measures.

Second, to improve equity in financing and access, it is important to have different cost sharing requirements for different people. In most cases, cost sharing in the United States and some European countries (e.g. Norway) is unrelated to income, and thus regressive. Individuals and households with lower incomes will spend a greater proportion of their income just to meet the cost sharing requirement. This problem is accentuated by the fact that lower-income people tend to be in poorer health and thus need more care. Although practically there might be difficulties in defining and identifying those who should be exempted from cost sharing, theoretically the rate of cost sharing should vary according to people's ability to pay.

Third, to improve allocative efficiency, it is important to have different cost sharing requirements for different types of health interventions. Evidence shows that the higher the cost sharing rate, the more reduction in utilization would be expected. Presumably, if cost sharing rates are higher for less cost-effective interventions and lower for more cost-effective interventions, the variation of the rates will have an effect on allocative efficiency. Although practically it might be difficult to define and identify the cost-effectiveness of many health interventions, theoretically the rate of cost sharing as a price signal will at least in part affect the choices of patients and direct the allocation of limited resources.

Chapter 5. Defining Benefit Packages

Defining a benefit package of health interventions as policy works from both ends of the list of possible interventions. A positive approach defines which interventions should be included in the guaranteed package. A negative approach defines which interventions should be excluded from a package. These two approaches are driven by two facts. One is that in low- and middle-income countries the most cost-effective health interventions are not always accessible to the population, especially the poor, and efforts must be made to find mechanisms to finance these interventions (12). The other fact is that in high-income countries the increase in the cost for publicly financed programmes makes it impossible for the government to guarantee all health interventions, and efforts must be made to exclude some interventions from the publicly financed packages (277). These two approaches are related to two policy tools: rationing and establishing essential packages of interventions.

Rationing

Rationing is defined as activities which involve denial or dilution of some health care that is potentially beneficial to the patient, so the patient is getting less in the way of treatment than might be thought desirable in a world with unlimited resources (278). As a policy tool, rationing can be simply defined as exclusion of some costly and cost-ineffective care from the benefit package of health care which is covered by social health insurance schemes.

Rationing was formally adopted only in the early 1990s, when Oregon State in the United States decided to cut down the benefit package of the Medicaid beneficiaries to allow for the expansion of coverage. Rationing had in fact existed in all health care systems across all countries, for example, through ability to pay, waiting lists for some sorts of services in publicly funded provisions, and rationing by medical doctors under a constrained government budget. Rationing is different from priority setting, first because the latter is a broader concept (one can prioritize groups of people, types of services, and types of diseases), and the former is related to denial of some types of services; second, the former affects individuals in terms of their access to some types of services, whereas the latter affects a group of people. Rationing is related to the activities for denial of the costly or less cost-effective services within the government-funded financing systems.

The rationale for rationing is straightforward. All governments are unable to guarantee all services for all people because of the limitation of resources. There are always choices among three variables: the number of people covered (the width of coverage), the types of services covered (depth of coverage), and the cost for coverage of the people and services. If a greater number of people are covered, costs will rise unless the number of covered services is cut; if the number of covered services is increased, costs will rise unless the number of covered people is reduced (279). If one of the major goals of the health care system is to maximize health for a given budget constraint, adoption of universal coverage for the more cost-effective services until the budget runs out becomes a reasonable strategy. Given the limitation of government budgets, rationing is recognized to be inevitable in a publicly financed health care system (278).

In the United States, the increase in the cost per Medicaid beneficiary and the constraint of government budget made many states lower eligibility standards for Medicaid to an income

level well below the federal poverty line, and the number of beneficiaries was cut down. As a result of this, fewer than 50% of the poor were covered by the programme (280). Reformers in Oregon promised an alternative to the practice of denying coverage to the insufficiently poor by extending Medicaid coverage to all persons living below the poverty line. The enlarged Medicaid enrolment was said to be financed by covering fewer services. In other words, expanded access to health insurance for the poor was to be pursued by rationing their medical care. In order to determine which services to fund for Medicaid beneficiaries, the Oregon Basic Health Services Act charged an 11-member Health Services Commission (HSC) to create a list of health services ranked from the most to the least important, representing the comparative benefits of each service to the entire population to be served. This commission reduced over 10 000 services to a prioritized list that initially ranked 709 condition and treatment pairs. The legislature's decision on how much to spend on Medicaid was presented as "drawing a line" in this list, financing only those services above the line (281). It represented a striking contrast both to the implicit rationing of medical care in the United States by income and insurance coverage, and to the somewhat less visible resource allocation decisions by health policy-makers in other countries, where some of the services are implicitly rationed through waiting.

The HSC collected cost and effectiveness data and ordered their first prioritized list entirely by cost-effectiveness ratios. This first list was unveiled in 1990 (and not then submitted to the federal government), but the ordering of services was considered by many observers to be deeply flawed. Rather than attempting to improve their data so as to continue working within the cost-effectiveness paradigm, the HCS abandoned this method and went on to adopt different ranking algorithms, producing revised lists in 1991 (which was rejected by the Bush administration because it violated the Americans with Disabilities Act), 1992 (which was conditionally accepted), and 1993 (which was approved by the Clinton administration) (281). For the 1991 list, the commission divided the 709 services into three categories: essential, very important, and valuable to certain individuals:

- Essential: services that preserve life, maternity care, preventive care for children and adults, reproductive services and comfort for the terminally ill.
- Very important: treatment for non-fatal conditions in which there is full or partial recovery and treatment that will improve the quality of life.
- Valuable to certain individuals: treatment for non-fatal conditions which merely speeds recovery and those in which treatment provides little improvement in quality of life.

The commission provided a cost estimate for providing various levels of service in the state Medicaid programme. If the third category was eliminated, there would be a saving of 22% of the cost, which could be used to expand the number of covered people (282). However, the later versions of this list considered other factors, such as prevention, quality of life, ability to function, equity, effectiveness of treatment, benefits for many, mental health and treatment for chemical dependency, public choice, community compassion, impact on society, length of life, and personal responsibility. These factors were taken into account by the members of HSC through ad hoc adjustment of the lists (283). As a result, the list became a product of a societal decision whereby many criteria mattered and decision-makers needed to consider various trade-offs.

Criticism existed throughout the process of rationing. According to Leichter, criticisms of the Oregon plan included: first, the medical value of a list that sought to cram more than 10 000 medical diagnoses into 700 or so lines containing condition/treatment pairs; second, the list did not only consider the cost-effectiveness criterion; third, the limitation of access to some health services only for the poorest and most politically vulnerable population; fourth, the potential

for a two-tier system, where doctors would have to be asked to practise two-tier medicine explicitly, in which patients with resources would receive some beneficial treatments and Medicaid patients would receive fewer treatments (284).

In the United Kingdom, services have been implicitly rationed through general practitioners (GPs) as gatekeepers and through the professional autonomy of hospital doctors. Both mechanisms are buffered by hospital waiting lists. Rationing for drugs has been implemented through GP fundholding and providing incentives to both providers and patients. To deter the use of expensive and relatively cost-ineffective drugs, GPs and GP fundholders are encouraged to increase generic prescriptions. GPs are given a budget for drugs, and the savings can be kept by GPs, provided that the quality of prescribing does not deteriorate. A co-payment rate of £5.75 for each item of prescription was reported to have an effect on reducing demand (252). The debate on explicit rationing was intensified in the early 1990s in the United Kingdom, and most writers concluded that rationing was inevitable, but explicit rationing as in Oregon seems to be impossible, given the history of NHS and the political context in the United Kingdom (285). In the United Kingdom, rationing is such a major political issue that the term rationing has become politically unacceptable. So far there has been little public debate on what should be excluded from public funding, although generation of cost-effective information on health interventions has become both a national and an international effort (e.g., the establishment of the National Institute of Clinical Excellence in the United Kingdom[2], and CHOICE[3] in WHO). It is recognized that it is more difficult to withdraw a service that has been already provided than not to introduce new services, for both political and economic reasons (252).

In Australia, there was a small-scale demonstration which attempted to develop explicit criteria for overt drug rationing (286). The government developed a formal scoring system to prioritize drug introduction. The guiding principle was to use the drugs so as to achieve the greatest benefit for the most patients for each dollar spent. The benefit measure consisted of three elements: outcome score, types of treatment, and clinical comparison with other treatments. The cost score also contained three elements; a cost comparison with alternative treatments, an assessment of the total cost implications underlying the introduction of the drug, and the cost per complete treatment course. While this approach provides some opportunity for rational, equitable and transparent judgement in restricting the drug budget, it was criticized as a formularized approach that was far from medical practice (287). This approach was not put into wide practice in Australia.

In the Netherlands, a committee was appointed in 1992 by the government to advise on the benefit package to be included in the social insurance system (288). This committee provided a framework to guide decision-making. In brief, the committee recommended that services should be required to pass four sequential sieves (tests) before they were included in the benefit package. The first sieve was necessity, which was defined as care that responds to basic need and is necessary to maintain or restore health. The second sieve was effectiveness, according to which only the care that is effective in maintaining or restoring health should be included. The third sieve was efficiency; so among the alternatives of effective care, only the most cost-effective are included. The fourth test wasindividual responsibility, which means that only the care that cannot be left to an individual's responsibility is included. These principles did not, however, lead to an explicit list of services.

[2] Web site: http://www.ncchta.org/
[3] The acronym CHOICE stands for "choosing interventions that are cost-effective".
Web site: http://www3.who.int/whosis/menu.cfm?path=evidence,cea&language=english

In New Zealand, the government established a committee in 1992 to advise on the core services to be funded within the National Health Service (289). At an early stage in its work, the committee concluded that it was neither helpful nor sensible to draw up a tightly defined list of services to be provided. Later on, the committee went on to define broad principles for priority services, and focused on the development of clinical guidelines for doctors to determine which patient should receive what type of care.

In Sweden, a committee was appointed by the government in 1993 (290). This committee analysed the issue of priority setting and worked out principles to guide decision-making. These principles included human dignity, need, solidarity, and efficiency. It was specified that services should not be rationed according to age and income. There is no report on a explicit list of services included or excluded.

Evidence of the effectiveness of rationing is limited only to Oregon State in the United States, because none of the other countries have used rationing as an explicit policy tool. It was originally projected by advocates of the Oregon plan that by not paying for costly and medically ineffective treatments (e.g., life-support services for extremely low-birth-weight babies or patients with brain deformities, and in vitro fertilization procedures), Medicaid expenditures in Oregon would fall by US$ 8.9 million or 22% (282), but in fact the Oregon Health Plan did not generate substantial savings. It was estimated that the list saved the state only 2% of total costs for the programme over its first five years of operation (280). The expansion of enrolment was largely financed through increases in funds from government general revenue, the imposition of a tobacco tax, and pushing more Medicaid recipients into managed care plans (33% of Medicaid beneficiaries were enrolled in managed care before 1993; 85% in 1997). The reason for this failure was that the health plan was not able to cut back the services, as originally proposed, for various reasons which are stated below.

According to the original idea, under the Oregon Health Plan the services covered would go from the top of the list towards the bottom until the budget ran out. The remaining services were to be excluded from public financing. In practice, however, the plan failed to produce a significant reduction in the total quantity of services. There are several reasons for this. First, full access to care by the poor was officially enshrined in law. Although the Oregon Health Plan was approved by the federal government, excluding services that were formerly covered is controversial and sometimes difficult to implement. Second, some services which had not been included before were included. For example, mental health services, which the government and the private sector had resisted, were not only included, but were subject to no limitations on the duration of care; HIV carriers was covered for full care because legislators were unwilling to cut off their coverage, which extended far beyond basic services (291). Third, the services that had been planned for exclusion were not always excluded. Doctors and hospitals provided services below the line that they considered necessary (5% of the total health care cost was actually for the below-the-line services). The reasons were the existence of co-morbidities between the covered and the uncovered services, and the failure to explicitly exclude transplantation from the service package.

Although it was originally planned that the choice of services would be based on cost-effectiveness, the final list responded to other considerations. Later lists were adjusted by the Health Service Commission, which considered public preferences and federal regulations and requirements. The reputed technical power of this administrative instrument had drawn many visitors to Oregon eager to study its precise formulas and innovative rationing tool.

Oregon did expand the number of people covered and their access to services. Oregon's Medicaid programme was able to cover all residents below the poverty line, and over 100 000 newly eligible Oregonians enrolled in the expanded Medicaid programme. While nationally the percentage of children without insurance rose from 14% in 1990 to 15% in 1996, the per centage in Oregon decreased from 21% to 8% in the same period (280).

The experience of Oregon and the efforts made in many other countries show that there will always be a trade-off between comprehensiveness and universality of coverage of health care. Although it is impossible to provide all health services to all people, it is much more difficult to exclude than to include services that have been included in a system in which there is a commitment to deliver comprehensive health care. Rationing is much more than a technical consideration; it is usually politically sensitive. International experience shows that there is a need to combine the use of rationing techniques with a process for drawing on the views of experts and the public. The most feasible approach might be as in New Zealand, where the principles of priority setting are worked out and the practice of rationing follows clinical guidelines for treatment. The innovation of Oregon is more political than technical. The real innovation in Oregon is the development of a coherent political strategy to accomplish reform in a national environment hostile to social reform and to put explicit rationing of health care on the agenda.

A number of issues need further debate by both health policy-makers and researchers. The first issue is the way in which health care is rationed. Alternatives include exclusion (the denial of some services) and dilution (making some services less accessible); explicit (providing a clear list of what services are financed and what are not) and implicit rationing (providing guidelines and incentives for medical providers to decide on the actual provision of services). Experience shows that exclusion and explicit rationing are extremely difficult in countries where a generous commitment to coverage has been made and people have enjoyed access to almost all types of services. The second issue is related to the criteria used for rationing, including effectiveness, costs, cost-effectiveness, the rule of rescue (by which dying people should be rescued regardless of costs), personal responsibility, and the difference between public goods and individual goods. While some think all of these should be considered in combination (283,292,293), the practical use and combination of these criteria in rationing still needs debate. The third issue is who should be responsible for rationing: government, organizations or medical doctors, or all of them? It seems that without clarification of these issues, rationing will be something to talk about, not something to practise.

Defining essential interventions

"Interventions" are usually defined as all services, activities and medical products that are used to improve the health of people. Essential interventions are divided into two categories, namely an essential package of services and an essential drug list, with the former including services and activities, and the latter including all other medical products. Essential packages of services and essential drug lists actually go hand in hand: drugs cannot be delivered without services, and the delivering of services usually includes drugs. But because these are usually regarded as two independent concepts in the literature, we discuss them separately.

Essential packages of services

An essential package of services is generally defined as a set of health services that are considered important and that should be accessible to everyone (294). There are three approaches for defining the essential package: first, specify a list of services that ought to be available under any form of health care system (the problem of this approach is that criteria for selection of the services are not specified, and the approach is not operational); second, look at medical benefits and seek to develop an essential package out of known efficacious treatments (the problem of this approach is that an essential package of services is defined independently of resources, and society can not bear the costs for all types of interventions which may have medical benefit); and third, determine what the average person would want with respect to medical care and use that as a standard for determining what is essential care (this approach also defines essential services independently of resources, and the care wanted may have no medical benefit) (295).

Callahan lists five priorities, the highest being those services that can cure disease. The second priority should be those measures that promote public health (immunization, screening, health promotion, and disease prevention). The third priority should be primary and emergency care. The fourth priority should be advanced technological medicine, including elaborate forms of surgery, chemotherapy, and the like. The fifth and final health priority should be highly advanced technology, particularly organ transplantation and kidney dialysis (295).

All these definitions and arguments fail to introduce economic concepts into the definition. An essential package of services cannot be defined unless at least the cost-effectiveness of services and the resources for providing these services are considered. In addition, the definition of an essential package of services is a practical issue. Without an explicit list of services and policies to assure access to them by all people in a society, the essential package of service will not be a policy tool.

The rationale for defining and implementing an essential package of services is related to the recognition that no country in the world can provide health services to meet all the possible needs of the population, and the observation that while the cost-effective services which deal with diseases and risk factors that bring about the greatest burden of disease are not adequately provided, substantial resources are used for less cost-effective services. In addition, because inadequate use of essential services is more likely among the poor, to define and make accessible essential services not only promotes efficiency of resources use, but also improves equity of health and health care.

The principal argument for a collection of services to be provided jointly in an integrated manner (rather than simply a list of interventions that are provided separately) is to minimize the total cost of the package by exploiting the shared use of inputs and by reducing the cost to society for patients obtaining services. The packaging of services can assist government to set explicit priorities which define the use of resources and improve efficiency. It also helps to define the responsibility of the government for financing health care services (296).

Two criteria were used for the definition of an essential package of services by the World Bank: the size of the burden caused by a particular disease, injury or risk factor; and the cost-effectiveness of interventions to deal with it (12). A minimum essential package of interventions was proposed to deal with 71% of DALYs lost for children under 15 years old, and 50% of the disease burden for adults (296). It was estimated that 10-18% of the burden for adults and

21–28% of the burden for children can be eliminated at a cost of $12 per DALY for low-income countries and $22 per DALY for middle-income countries. It was suggested that for very poor countries, priority should be given to public financing of health care for the poorest of the poor and the most cost-effective interventions of the package; and for middle-income countries the government should finance all interventions in the minimum package for both the poor and the non-poor.

Following the advocacy of the World Bank, several countries tried to design an essential package of interventions. These countries include Bangladesh, Colombia, Eritrea, Ethiopia, Ghana, India, Indonesia, Kenya, Mauritania, Mauritius, Mexico, Turkey, Uganda, United Republic of Tanzania, Uruguay, and Zambia (294). In practice, however, few of these countries have yet implemented an essential package. Tarimo provided examples for four countries: Bangladesh, Colombia, Mexico and Zambia (294). Colombia and Mexico have developed a basic package of care to cover the low-income population that was excluded from formal health insurance. It aims to cover low-cost, high-impact interventions with an emphasis on health promotion and preventive measures. It is designed to build on the experience of existing insurance and provision programmes and to use a horizontally integrated approach. The package is seen as the irreducible minimum that must be provided nationally, but there is scope at regional level for added services. The interventions included in the package are basic sanitation for the family, effective management of diarrhoea in the home, anti-parasitic treatment, management of acute respiratory infection, prevention and control of TB, prevention and control of diabetes and hypertension, immunization, monitoring of child growth and nutrition, family planning services, services related to childbirth, prevention of accidents, and initial management of injuries. These services are planned to be provided by all levels of the health system (community level, primary level, secondary and tertiary levels) with variations in their responsibilities. The packages are intended to be financed by government.

Bangladesh and Zambia have worked out an essential package which includes an expanded programme on immunization, prevention and control of acute respiratory infections, diarrhoea, malaria and measles, maternal health, family planning, prevention and control of STD/HIV, prevention and control of TB and tropical diseases, prevention and treatment of malnutrition, treatment of common conditions, and medical emergencies. As with Colombia and Mexico, the responsibility of provision is planned to be shared by all levels of the health care system. The operational definition of the package is financially and technically supported by the World Bank, the World Health Organization and other international agencies, but there is no information on the implementation (financing and provision) of this package.

In Indonesia, the government mandates that health insurance schemes must cover a defined benefit package which includes the most cost-effective inpatient and outpatient services. However, only around 12–15% of the population are covered by health insurance, so this benefit package is not ensured for the majority of the population. The World Bank's fourth health project in Indonesia, running from 1996 to 2001, had as one objective the improvement of resource allocation by allocating a greater share of government health spending to the essential package of services. But there is no published information on the effectiveness of its implementation.

The merit of the essential package of interventions is that priority-setting is explicit. The package can guide the decision-making process in resource allocation and direct the resources to the most cost-effective interventions, because it provides operational information on what should be

financed and provided. Unfortunately, the process often stops at the stage of definition, and implementation is usually difficult. There seem to be no publications on the details of implementation and there is no empirical evidence on the extent to which the implementation of such a package has led to an improvement in allocative efficiency. The introduction of essential packages has initiated thinking about cost-effectiveness, but evidence of its use in guiding the reallocation of resources is so far limited (1).

The concept of an essential package of interventions has been accepted by policy-makers and researchers, but there are few cases where the policy is implemented, probably because of the following problems (294). First, there is a lack of guaranteed and sustainable sources of financing. In some poor countries there seem to be insufficient resources to finance the package of interventions. A study in Uganda estimated that to provide the essential package of interventions would require 56% of total health expenditure. A study in Zambia showed that the cost of implementing the defined package far outweighed the government/donor resources allocated to district level. Experience shows that sufficient and sustainable financing is the most important factor that determines the successful implementation of the package. Second, the willingness to implement an essential package and its success depend on the strength of government policy. The usual case in low- and middle-income countries is that the government directs public resources more to the needs of a powerful elite, rather than adopting the implementation of the package as a system-wide policy. As a result, the defined package may benefit the rich rather than the poor. Third, there exists a technical problem. Because of limitations of data and technical capacity in a particular country, the package is often determined by estimates about the effectiveness of interventions and social judgement about reasonableness, rather than by evidence that results from systematic and scientific investigation. The outcome is that resources may not be reallocated in an economically desirable way. Last, the lack of a capable and sustainable provision system and quality assurance programme may hinder the implementation and effectiveness of the essential package.

Essential drug lists

Essential drugs are defined by the World Health Organization as those that satisfy the heath care needs of the majority of the population and that therefore should be available at all times in adequate amounts and in the appropriate dosage and forms (297). The criteria of essential drugs include established efficacy and safety, coverage of the majority of the diseases prevalent in the population, adequate quality, and stability. The selection of essential drugs also considers the total cost of treatment, changes in public health priority, and levels of health facilities.

The main objective of the essential drug list is to increase the availability and accessibility of cost-effective drugs to the populations whose basic health needs have not been met by the existing supply system. A survey conducted by the World Health Organization showed that 70% of the population in 23% of the low- and middle-income countries had no access to essential drugs (298), reflecting the rationale for the establishment and implementation of such a list. The second objective, which is less explicitly stated, is to decrease the use of less cost-effective drugs.

In 1975, the World Health Assembly requested the World Health Organization to advise Member States on the selection and procurement at reasonable costs of essential drugs of established quality, corresponding to their national health needs. An expert committee proposed general principles and guidelines for establishing a list of essential drugs, and then

began to work on it. For much of the 1980s, the pharmaceutical industry resisted the spread of the essential drugs approach to high-income countries and to the private sector, but because of the well-accepted facts of cost escalation and inefficiency in drug utilization, the resistance is coming to an end. The World Health Organization expert committee has continued the work and produces and updates a model list of essential drugs every 2 or 3 years (299).

Since the first model essential drug list in 1977 (which included 205 drugs), 10 model lists have been provided. The latest model list includes about 300 drugs (297). Following the advice of WHO, some 110 countries have developed their essential drug lists (215,300). At the national level the list usually consists of 200 to 400 essential drugs. Many countries have worked out the essential drug list for different facility levels. For example, in Uganda there are 29 essential drugs for rural health centres and 277 for hospitals; in Viet Nam, there are 31 essential drugs for community clinics, while the national list includes 120 drugs (301). The concepts of essential drugs and model lists have also been used by many high-income countries to create national formularies to promote rational use of drugs and to reduce health care costs (297).

The essential drug list was particularly used in the Bamako Initiative in Africa, where essential drugs were hardly available because of the deterioration of publicly funded health systems. The Initiative was sponsored by WHO and UNICEF and announced by African Ministers of Health in 1987. It aimed at universal accessibility to primary health care (302). As an important part of the initial formulation of the Initiative, the donor agencies and governments provided revolving funds for community pharmacies or health centres to purchase essential drugs from central procurement organizations at lower than market prices; drugs were to sell at two or three times the wholesale price to cover the operating costs (214). During the 1990s, many African countries (302-306) and some South-East Asian countries (307) implemented such a scheme.

There is some evidence that the implementation of the essential drug list improved availability of essential drugs, and decreased the production and utilization of drugs of doubtful effectiveness, but the implementation of the programme varied among countries. In Bangladesh, since the adoption of the essential drug list, 300 essential drugs have been approved; 1600 products have been proved ineffective or harmful and their production has been banned. After the introduction of the programme, the local production of essential drugs increased by more than 217% from 1992 to 1999 (301). The implementation of essential drug lists as a part of the Bamako Initiative was reported effective in increasing availability and accessibility (303-305,307), although problems of inequity and over-prescription and commercialisation are increasing concerns (302,306).

To assess the availability and rational use of drugs, Hogerzeil et al. compared a random sample of 19 peripheral health units in Yemen, in which an essential drugs programme had been operational for the past few years, with seven health units in which no such programme had been started (308). It was found that the programme had significantly improved the availability and rational use of essential drugs in peripheral health units.

According to the literature review by le Grand, a randomized control study in Yemen showed that after the introduction of an essential drug list, doctors' knowledge of the rational use of drugs increased significantly; improvement of drug prescribing behaviour was observed: there was an increase in use of essential drugs, a reduction in injections, and a decrease in inappropriate use of antibiotics. Likewise, in Sudan, significant improvements were noted in the use of essential drugs in all health facilities after introduction of the essential drug list. In Ethiopia, the introduction of the essential drug list resulted in a significant decrease of non-essential drug prescriptions. Introduction of an essential drug list is most effective if

accompanied by an introductory campaign and adequate follow-up, and in countries with strict registration and regulation of drugs (215).

In some countries, the implementation of the essential drug list showed a less effective outcome. In Thailand, the essential drugs list includes 372 drugs from among over 20 000 types of drugs on the market. The government required that 80% of the drug budget in the Ministry of Health hospitals must be used to purchase essential drugs, and 60% by other public hospitals. In practice, however, the value of such drugs purchased by the government was only 37–48% of the total drug budget (301). In Nigeria, the World Bank provided a loan of US$ 68 million for an essential drugs project, aimed at increasing the access for cost-effective drugs. The report noted that the management of procurement and transportation of drugs was hampered by poor infrastructure and by corruption, and the inefficiency adversely affected procurement activities: the private sector made up 80% of the pharmaceutical market, and the drugs available on the market (mostly produced locally) were of largely unknown safety and efficacy (301).

Chapter 6. Informing Providers

In the past two decades, researchers and policy-makers have recognized that the utilization of health care and its costs are to a large extent determined by health care providers. The provision of cost-effective health interventions depends on two major factors: providers must have the best information on health interventions, and they must be motivated to use the information to provide cost-effective care. The latter has been addressed in Section 3, on economic incentives for providers. This section will discuss the former, that is the tools related to informing providers, in order to influence their recommendations for health care interventions. The tools to be discussed in this section include medical practice guidelines, prescription formularies, and utilization review.

Medical practice guidelines

Medical practice guidelines are defined as systematically developed statements to assist practitioner and patient decisions about appropriate health care for specific clinical circumstances (309). The guidelines usually offer concise instructions on which diagnostic or screening tests to order, what medical and surgical services to provide under what conditions, how long patients should stay in hospital, or other details of clinical practice.

Medical practice guidelines (sometimes called clinical guidelines, medical guidelines, treatment protocols, treatment parameters, or medical pathways) originate practically from utilization review and are theoretically related to evidence-based medicine, because review of the medical practice of doctors needs prepared guidelines, and the guidelines should be based on scientific clinical evidence. Thus medical practice guidelines are the practical form of evidence-based medicine, which is defined as the conscientious, explicit and judicious use of current best evidence in making decisions about the care of individual patients.

The implementation of such guidelines is justified by the fact that there is a large variation in medical practice for the same health conditions across geographical areas and among physicians. This suggests the inappropriate provision of health care which may cause unnecessary increases in health care expenditure without bringing about improvements in health. Wennberg et al. have argued that exposing clinicians to scientific evidence about medical interventions and information on the extent and effects of variations will reduce variations and improve providers' behaviour, and thus improve the cost-effectiveness of health care (310). Lexchin has argued that physicians' knowledge base is important for their decisions in choosing interventions for their patients. Biased information provided by market promoters (e.g., sale representatives of the pharmaceutical companies) may lead to distorted provision; and information on the cost-effectiveness of health interventions will improve medical practice and thus allocative efficiency (78).

Perceptions among public policy-makers that there is an inappropriate increase in health care costs, and that some medical care is unnecessary, have induced the medical research and practice communities to work together to develop practice guidelines. Well-designed medical practice guidelines are believed to help ensure appropriate utilization of medical services and improve the quality of patient care. They can also provide the foundation for developing

rational review criteria for quality assurance programmes, utilization review systems, and insurance coverage policies (311).

Evidence-based medicine and practice guidelines have been embraced by increasing numbers of scholars, administrators, and medical journalists as an intellectually attractive solution to the dilemma of improving health care quality while reducing costs. As Larson argued, although medical practice guidelines alone are not a solution to the problems of increasing costs, they can be an important input for the improvement of health care quality and the reduction of unnecessary medical costs (312).

Medical guidelines are increasingly used in various countries. In the United Kingdom, guidelines based on consensus conferences and expert opinions have existed for decades, but the rigorous design of guidelines based on scientific evidence is a phenomenon of recent years. Professional bodies produce guidelines to be used by providers to improve quality of care and by purchasers to guide contracting. In the Netherlands, the Dutch College of General Practice has produced guidelines since 1987. At present, more than 70 rigorously developed guidelines are used by medical practitioners. Finland has issued more than 700 guidelines since 1989. A programme of evidence-based guidelines development has been started recently. In France, there are 100 guidelines based on consensus conferences or modified from guidelines in other countries. In New Zealand, the implementation of guidelines is a national policy. Guidelines are intensively used to restrict less cost-effective services at the point of services.

In the United States, the idea of practice guidelines is not new. The American Academy of Pediatrics was the first organization to develop parameters, more than 50 years ago. By 1980, eight medical societies in the United States were developing practice guidelines of their own. In 1990, more than 26 physician organizations had developed guidelines, and approximately a dozen more societies had plans to do so. In total, these societies have produced more than 700 practice guidelines. The Clinical Appropriateness Initiative (a collaborative project between the RAND Corporation, the American Medical Association, and the Academic Medical Center Consortium) was one of the largest medical guidelines development efforts in the United States (311). The development of guidelines has spread from utilization review agencies to physician speciality societies, federal agencies, managed care organizations, and academic health centres, as their usefulness in accomplishing the goals of health care reform became clear (313). Although guidelines are developed by government panels and professional bodies, many health care organizations purchase commercially produced guidelines in the interest of reducing cost.

Evidence on the effect of medical practice guidelines generally supports the hypothesis that their use can improve quality of care and reduce health care costs. In France, surveys suggested that 75% of French doctors were prescribing in line with the new treatment guidelines, and the introduction of the guidelines was associated with a 15% reduction in prescribing of antibiotics in the first six months of 1994; prescriptions of non-steroid anti-inflammatory drugs and anti-ulcer drugs were also reduced. The authors argued that apparently these were the results of the guidelines (314). A literature review on the effectiveness of guidelines on periodontic treatment showed that they minimized treatment variability and ensured the quality and cost-effectiveness of care by improving health outcomes and reducing costs (315). An intervention study found that using computer-assisted prescription guidelines could affect physicians' prescription behaviour: antibiotic use decreased by 22.8% overall, and clinical health outcome improved, and these results meant a considerable improvement in efficiency (316).

Grimshaw et al. reviewed 59 published evaluations of clinical guidelines that employed either randomized control or before-and-after comparison with a control group (317). Of these

studies, 24 investigated guidelines for specific clinical conditions; 27 studied preventive care; and 8 looked at guidelines for prescribing or for support services. All but four of these studies detected significant improvement in the delivery of care after the introduction of guidelines; and all but 2 of the 11 that assessed the outcome of care reported significant improvement. The authors concluded that explicit guidelines do improve clinical practice and health outcome, but the size of the improvement in performance varied.

While it is generally accepted that best-practice guidelines will improve clinical performance, there may be a bias to report positive results, as Goldfarb argued. The effectiveness of the use of guidelines depends on how good they are and on the actual usage of the guidelines in medical practice. In addition, because the development and implementation of guidelines involve costs, whether or not the use of medical guidelines can reduce overall costs needs continuous research (318). The development of medical guidelines has mainly been the result of government efforts in Europe, but private insurance companies in the United States are also involved. The likelihood of biased medical guidelines being developed by profit-driven insurers has not been examined.

Prescription formulary

A formulary is a list of pharmaceuticals permissible to use in a health insurance programme. It is continually revised reflecting the current judgement of medical professionals (319). There are three types of formularies: open, closed, and preferred. An open formulary serves merely as a guide, including a list of drugs for which relative cost information is provided. A physician may prescribe any drug, but is encouraged to use the formulary list in prescribing decisions. In contrast, a closed formulary lists the drugs that will be reimbursed by third party payers; non-formulary drugs will be reimbursed only if they are authorized prior to prescribing. A preferred formulary imposes lower co-payments for drugs in the formulary and higher co-payment rates for drugs falling outside the formulary (319,320).

The types of drugs included in the formulary are decided by a committee including physicians and pharmacists, based on efficacy, safety, and cost-effectiveness information provided by the governmental drug administration, pharmaceutical companies and published materials. A formulary lists drugs by therapeutic class and provides relative cost indices. The use of cost-effectiveness information in the composition of the formulary is still limited, first because comparative cost-effectiveness information between or among substitutes is hardly available, and second because the inconsistency among researchers in measuring costs and benefits makes the information hardly comparable (320).

The stated objective of using formularies is to reduce costs for drugs and increase quality of care by reducing inappropriate use. The formulary system has, however, been severely criticized for putting too much emphasis on cost and too little on quality of prescriptions. Moreover, a common complaint is that formularies often include only those drugs that are cost-effective for the average patient, while overlooking the special needs of individuals (319).

Despite these criticisms, formularies are intensively used in many countries of the European Union, where the health systems follow a pattern of NHS or social health insurance schemes (77,321). In the United States, the traditional indemnity plans do not include drugs as a part of

the plan's benefit, but in managed care organizations (e.g., HMO) drugs are increasingly included in the benefit package. Formularies are used not only in the Medicaid programme, but by most private health insurance plans as well.

Although the use of a formulary is expected to increase the quality of prescribing as well as reduce the costs of drugs, research findings have suggested that the implementation of formularies can only reduce costs for drugs. Whether it can reduce overall cost is highly controversial. In South Africa, a study analysing the gap between prescribing based on a formulary and current prescribing in a large HMO found that the use of the formulary can reduce drug costs by 20% (322). The implementation of a formulary in all hospitals in Hong Kong Special Administrative Region, China, was reported to have rejected 13 cost ineffective drugs and saved about HK$ 191,000 from April 1995 to February 1996 (323). A literature review of studies in the United States showed evidence that well-controlled formularies could reduce drug expenditure in hospital settings (324). Another extensive literature review of studies in the United States found no evidence to date that the use of formularies adversely affected patient access to necessary pharmaceutical care, or worsened health outcomes (325).

Many studies show that physicians' prescription behaviour is affected not only by the costs of drugs, but also by who pays for them. In a mail survey in Canada, 1070 primary care physicians were provided with the clinical scenario of a patient with an infectious exacerbation of chronic obstructive pulmonary disease, and asked to select diagnostic tests as well as one of six antibiotics. Two antibiotics were expensive (ciprofloxacin and cefaclor; average price $Can 52.23) and four were inexpensive drugs (amoxicillin, Cotrimoxazole, erythromycin and tetracycline; average price $Can 2.8). Neither of the two expensive drugs is considered a first-line therapy for the condition described. Questionnaires differed in the presence or absence of drug benefit coverage and price information. With third party cost coverage and the prices shown, 18% of respondents selected an expensive antibiotic. This increased to 38% when the prices were omitted and decreased to 8% when the patient was said to have no drug benefit coverage (326). Several other studies also showed that physicians were sensitive to the costs of drugs, and more sensitive to the cost for the patient than for the third party (261,327-330). These studies provide plenty of evidence that formularies that includeinformation on cost, effectiveness, and who bears the costs affect prescriptions.

The question that remains unanswered is whether a formulary will reduce overall health care costs (331). There is substantial evidence that the use of formularies can reduce costs for drugs, but the reduction is achieved along with an increase in costs for other services. Horn et al. studied the relation between formulary limitations and severity-adjusted resource utilization for 12 997 patients with five tracer diseases in six HMOs (332). They found that the degree of limitation of the formulary was associated with higher resource utilization, measured by the number of physician visits, the days of hospitalization, and the number of prescriptions. They concluded that the restriction on the selection of drugs had shifted costs to other services and led to the overall increase in health care costs. Horn claimed that at least 30 other studies support this study's findings (333).

A literature review by Kozma et al. concluded that if a formulary is implemented for expenditure reduction only, and many first and second choice drugs are not available, programme expenditures may actually rise because of the use of other more expensive programme services (e.g., hospitalization); reductions in restrictions of the formularies have no effect on drug programme expenditure (328). The study of Medicaid patients in South Carolina using a large sample (12 000 patients) by the same authors showed that following the

reduction in formulary restriction to Medicaid patients, there was a small increase in drug expenditure, but they also observed a reduction in hospital utilization and a decrease in costs for other services. The findings support the argument that health interventions are interrelated, and strategies that focus only one type of intervention (e.g., drugs) may increase the cost for others.

Despite the widespread use of prescription formularies, the impact of formularies on the quality of care for patients has not been well studied. A few controversial studies have indicated that formularies led to either higher costs or lower quality of care for some patients (332). It is argued that much work remains to be done before any sound conclusions can be reached.

Utilization review

Utilization review (UR) is defined as evaluation of the necessity, appropriateness, and efficiency of the use of medical services and products (334): that is, whether the right interventions are provided to the right patients, at the right location, for the right period of time. Although UR could be used to detect both underutilization and overutilization, in practice its focus has been the latter.

UR can be internal or external. Internal UR is usually carried out by a health care provider's own staff. When a problem case is found, the responsible doctor is usually provided with feedback, and asked for explanations and possible corrections. Thus internal UR is less punitive and more collegial. External review is carried out by UR agencies outside the organization where medical services are provided. External review is often more impersonal. If a problem is detected, the responsible doctor may get feedback or penalties including denial of payment, complaints to licensing bodies, discontinuation of the contract between the third party and the provider, and even criminal charges.

UR can be prospective, concurrent, or retrospective. Prospective review occurs before services are provided. This review is often carried out for elective admissions, costly drugs, and major procedures, such as surgery. The purpose is to get permission for provision, so that the payment for the delivered services can be guaranteed. Concurrent review takes place after admission but before discharge, particularly if the patient needs to stay much longer than the standard length of stay.Retrospective review occurs after discharge, when the entire procedure can be reviewed.

UR can be divided into three levels. First-level review is a screening process in which cases are judged by quantitative criteria or objective norms. This level of review is often conducted by trained personnel (not necessarily nurses and doctors). Suspicious cases (e.g., those with longer stay than considered appropriate) will be selected for the second level of review, which is often conducted by registered nurses with extensive clinical experience.The reviewers make judgements based on information gathered from patients' charts, telephone calls to the patients, interviews with doctors, their own experience and medical guidelines. The reviewed cases can be certified as appropriately handled, denied (with the possibility of appeal), or referred to the third-level review. The third-level review is done by physicians with substantial clinical experience. Physician reviewers make the final decision regardless of criteria.

UR initially focused on hospitals and their services, but it is extending to all types of health care providers and medical products such as drugs. Reviews can be conducted by providers

themselves (e.g., hospitals), third party payers, government-sponsored agencies (such as Peer Review Organizations in the United States), and private companies that sell UR as a service.

The force behind the development of UR in medicine as a whole has been the concern with the rising cost of health care. Many have feared that a significant part of the increase in health care costs is a result of unnecessary or inappropriate services and drugs. Thus the goal of UR is to improve efficiency (335), through reduction of costs and improvement of quality of care (336).

Theoretically, UR can improve resource utilization in several ways. First, the denial of services following prospective and concurrent reviews to a large extent means that the funds will not be spent. Second, although the denial of payment to providers by the third party does not mean saving of resources, this penalty will keep providers aware that UR is taking place and force them to take actions to reduce unnecessary provision. Third, UR can help to increase the knowledge of providers increasing the likelihood of cost-effective practices.

The country where UR is the most popular is the United States; although UR as a quality assurance measure can be traced back to the earlier part of the 20th century, it did not become widespread until cost escalation of health care became a major concern. By the early 1970s, concern over the rapid escalation in health care expenditure for the Medicare and Medicaid programmes incited the Congress to create and fund the Professional Standard Review Organization (PSRO) to perform UR on hospitalized beneficiaries (337). Because of the evidence that PSRO has little effect on cost control, and conflicts between government agencies and medical professional organizations over professional autonomy and health care quality, the PSRO was replaced by the Utilization and Quality Control Peer Review Organization (PRO) in the middle of the 1980s, focusing more on quality of care. The Health Care Financing Administration in the United States contracted with state PROs to assure the quality of services and eliminate unreasonable and inappropriate care provided to Medicare beneficiaries. According to the law each state had to select 10 major procedures for pre-certification (338).

The most popular tool used in the United States is the Appropriateness Evaluation Protocol (AEP), which was developed by Gertman & Restuccia (339), who worked on this first at the University of California and later at Boston University. This tool has been tested in Israel, Italy, Spain and the United States with satisfactory results. Along with the development of managed care organizations, which depend on their ability to create a mix of incentives and external constraints that eliminate inefficiencies in clinical practice while not harming quality of care, the UR is increasingly being used by the prepaid health plans. Virtually all of the 60 million enrolees in prepaid health plans were covered by some form of UR (340).

Drug Utilization Review (DUR) forms a major component of UR in the United States. Large- scale DUR began in 1990, following the mandate of Congress, which required Medicaid programmes in all states to undertake DUR. The purpose is to eliminate unnecessary and inappropriate drug utilization, and thus to reduce drug-induced hospitalizations, improve quality of care, and reduce costs for drugs and hospitalizations. DUR is conducted by contracted agencies based on pre-set drug prescription criteria. If drug therapy problems are identified based on whether the physicians' prescriptions fit an "exceptional drug profile", they are examined in detail by peer reviewers. If a problem is identified, the responsible physician will be notified and asked for correction or explanation. Follow-up review will be conducted for the physicians with drug therapy problems.

Utilization reviews for hospital services have been used in European countries. In 1993, a group of researchers financed by a European Union grant reviewed the use of UR in Europe. Different expert groups from seven European countries participated in this process (341). The researchers found that UR had avoided assessing the appropriateness of surgical procedures and diagnostic tests in order to concentrate on the patterns of provision of hospital services and drug prescriptions. UR was introduced for various countries (Australia, France, Italy, Portugal, Spain, Switzerland and the United Kingdom), some of which started as early as the 1950s (e.g., the United Kingdom) and some as late as the 1990s. According to the review by Lorenzo et al. none of these countries conducted UR on a routine basis as was the case in the United States (341).

In contrast to UR of hospital services, DUR is very popular in Europe. In the United Kingdom, the prescribing analysis and cost scheme disseminates information about prescribing behaviour to general practitioners in the hope that it will increase their awareness of inappropriate prescriptions and costs. Prescriptions are collated by a national authority, and information is fed back to general practitioners either in simple headline format or in more detail when a practitioner's prescription cost is substantially greater than the local average.

Several other countries (France, Germany and New Zealand) have information feedback systems for physicians similar to the UK scheme. However, all of these strategies are advisory and provide information on the volume of prescribing and on costs, but they do not give information on the cost-effectiveness of prescribing and so may penalize the use of expensive drugs for which the benefits are worth the extra cost.

In the Netherlands, monitoring and feedback of medical care utilization and costs by sickness funds to health care providers are conducted routinely (342). Data related to utilization pattern (drug prescription, specialist referrals, the number of admissions, length of stay, surgeries, laboratory tests and radiological services) and costs for general practitioners, specialist, hospitals, physical therapists and pharmacies are gathered and analysed. In addition, there are regulations that limit the quantity of drugs that can be prescribed under the reimbursement scheme, with the aim of preventing waste and excessive use of drugs through high quantity prescriptions. If a doctor's prescriptions exceed the upper limit, the excess will not be reimbursed (314).

As regards evidence of the effect of UR, much of the literature is from the United States and focuses on hospital services. Wichizer's paper represents the most thorough review and it is divided into three parts: a review of studies of government-sponsored UR programmes before 1985 (e.g., for hospital services used by Medicare and Medicaid beneficiaries); studies of private utilization programmes before 1985 (e.g., UR conducted by private insurers); and a review of the literature from 1986 to 1990. For government-sponsored UR, he found that the 12 studies he reviewed presented a range of findings. Some studies found little or no effect, while others found that UR reduced hospital admissions by 10–15%, and average length of stay by 5–10%. Based on the analysis of the methods of each study used, he concluded that taken as a group the studies suggest that UR did reduce hospital use. Although it is difficult to state with precision what the magnitude of this reduction is, a reasonable conclusion would seem to be that UR reduces hospital inpatient days by somewhere between 5% and 10%. However, these studies did not address the effect of UR on costs for hospital services. In his summary of the review of the studies of the private UR programmes, Wichizer states:

> *"Studies of private UR programs have produced encouraging results showing that UR does reduce hospital use and constrain costs. The problem with these studies is that the validity of their finding is suspect because of methodological problems. Additional problems arise because the studies do not provide adequate descriptions of the UR programmes, making it difficult to interpret and compare finding across studies*".

In an effort to update the findings, he reviewed a number of more rigorous studies. All these showed a decrease in both use and expenditure. In general, UR reduced hospital admission by 12–13%, inpatient days by 8–10%, hospital routine expenditure (including room and board) by 6–17%, hospital ancillary expenditure by 9–14%, and total medical expenditure including both inpatient and outpatient services by 9–14%. The estimates reported by these studies apparently represented real effects (343).

While the effect of UR on hospital utilization and expenditure seems to be clear, there is no evidence on how UR affects the quality of care and patients' health outcomes. In addition, whether the UR and related financial incentives have led to reduction of unnecessary care is unclear. A report showed that the pre-certification programmes operated by the Heath Care Financing Administration's Utilization and Quality Control Peer Review Organisation (PRO) were unsuccessful in reducing unnecessary procedures.These programmes expended significant resources but found less than 0.2% of the procedures they reviewed were unnecessary. These programmes failed in part because they lacked comprehensive and valid criteria on which to base their decisions as well as a rigorous structure to implement them (344).

Because of the debate over the effect of UR on quality, the United States Congress called for the National Academy of Science to propose strategies for quality review and assurance in Medicare. The research report of National Academy of Science led to a further evolution of UR in the early 1990s. The original dual functions of PRO (quality assurance and cost control) were replaced by a new function of quality assurance only. Whether the change in function has led to improvement in health care quality is unclear, because of the lack of tools to measure health care quality and constraints on the quality of administrative data (345).

The effect of drug utilization review (DUR) in the United States seems to be controversial. According to several publications, DUR programmes resulted in about a 3% reduction in drug cost (346-351). These reports were, however, criticized by Moore for several problems. First, these studies were conducted by programme administrators or private consulting firms that had a vested interest in showing their programmes to be successful. Second, most studies lacked a control group, were statistically flawed, and exaggerated the savings. Third, DUR often went hand in hand with other policy interventions, such as prior authorization, the use of formularies, and co-payment for drugs. It was very difficult to differentiate the true effect of DUR from others. Moore argued that the most rigorous study, by SRI International (352), had concluded that DUR failed to yield any overall cost saving; and evidence in that study suggested that DUR might increase the cost (because of high administrative costs) and decrease the quality of care for some subgroups of patients as a result of the possible underuse of cost-effective drugs (353).

Evidence on DUR in other countries is generally positive. Bochner et al. reviewed DUR in hospital settings, and concluded that DUR programmes had been shown not only to improve the standard of medical care, but to result in substantial cost savings (354). In the Netherlands, the information related to high use and costs generated by UR was used to provide feedback for medical doctors. Efforts conducted from 1981 to 1985 reduced the overall referral rate by about 10%. Feedback to general practitioners on their referral and prescription practice had led to about a 14% reduction in costs (342). Two controlled studies in Mexico showed that peer review and feedback of prescribing patterns had yielded positive results (215).

Greco et al. nevertheless argued that although detailed DUR and feedback involved monitoring a physician's prescribing practices and then providing feedback along with specific recommendations for changes in prescribing, not all DUR and feedback programmes had been successful, suggesting that certain conditions must be met before physicians respond to the feedback (355). Bochner et al. suggested that the effectiveness of DUR depends on the way the prescribers are informed of the results of the review. A review of the literature shows that face-to-face interaction with the prescribers is the most effective intervention in reducing the quantity and costs for drugs (356). A randomized controlled trial conducted in Australia showed that feedback in the form of mailed educational newsletters had no effect on prescribing behaviour of GPs (357).

A literature review of DUR and feedback studies shows that the effectiveness depends on the method of feedback (78). The following feedback methods are not effective: large group lectures and conferences (because of lack of information to facilitate practice changes), mailing of printed materials and guidelines (because some guidelines are not practical, and there is no financial incentive to change prescription behaviour, besides the fear of malpractice litigation resulting from underuse of drugs). The most effective intervention is academic detailing — face-to-face interaction between detailer and physician to deliver educational messages on appropriate prescription. It was shown that effects from a 15 minutes detail could persist for two years and resulted in a 12–49% reduction in inappropriate prescriptions (358,359).

While evidence suggests that DUR and the use of appropriate feedback methods may reduce quantity and costs for drugs, there is a virtual absence of research on patient outcomes. Intuitively, one would expect more appropriate prescribing and less inappropriate use of drugs to produce better patient outcomes, but this question has rarely been addressed in the literature (78,342,356).

Practice profiling is increasingly used, but in contrast to utilization review, the effects are not well studied. Medical practice profiling is defined as the analysis of rates of events pertaining to the process and outcome of medical care provided by health care practitioners to a defined population. Rates may refer to monetary spending, number of services provided, or number of outcome events occurring in a given unit of time. Health care practitioners include the entire spectrum of professionals who make patient management decisions, such as physicians, dentists, physician assistants, and nurse practitioners. The population may comprise individuals defined by their use of specific providers, their eligibility for specific benefits or their residence in specific localities. The overall objective of profiling is to use epidemiological methods to describe medical practices, monitor health outcomes, and assess the efficiency and quality of care. Profiling can provide heath care managers, purchasers, consumers, regulators and policy-makers with information to compare providers on dimensions of cost, utilization, quality and access, detect potential problems related to overprovision and underprovision, cost, and quality of care, and identify providers with deviant provision decisions. Profiling can also be used to assess physician performance for purposes such as contracting a provider, credentialling, payment, and identification of underprovision and overprovision. Profiling can describe overall patterns of resource use and suggest areas for improvement in efficiency or quality of care. Valid profiling analysis depends on the availability of pertinent data, which are usually based on insurance claims and other automatic information systems (360).

Four features of profiling convey useful information about medical practice.

- First, profiling focuses on patterns of practice rather than individual instances of care; thus its coverage is comprehensive in terms of patients, providers, and types of heath interventions, so it is less costly than case-by-case utilization review.
- Second, profiling is less intrusive because it uses available administrative and claims databases, rather than reviewing the practice records of individual practitioners. Thus it is more acceptable and less resisted by medical professionals.
- Third, profiling makes possible the identification of systematic practice problems by comparing the practice of different doctors: for the same type of disease; with and without medical guidelines; and between geographical areas.
- Finally, profiling allows for a combination of profiling and utilization review, by which doctors with consistently abnormal practice will be examined in detail. If there is a problem they will receive feedback or penalties.

Because of the advantages of practice profiling over utilization review, the former was promoted in the 1990s. Detailed evaluation of practice profiling remains to be done. Theoretically, the success of practice profiling depends on three factors.

- First, the availability of electronic information systems, such as those for insurance claims, drug prescriptions, orders for tests and examinations, and patient medical records.
- Second, data of good quality (accuracy and representativeness) is of key importance.
- Third, the effects of profiling depend on how the information generated from profiling is used. Profiling will not be effective unless the information is used for improving providers' behaviour, through feedback, contracts, performance assessment, payment, and benchmarking medical practice.

Detecting medical errors has long been used to gather evidence for malpractice lawsuits and internal or institutional quality assurance programmes. Systematic efforts for detecting, reporting and reducing medical errors have not been made at country or regional level. Only recently, the United States government decided to build up a national mandatory medical error reporting system, and to search for methods to reduce medical errors (361). Specific methods and their effectiveness are still awaiting evaluation.

Medical error is defined as "the failure of a planned action to be completed as intended or the use of a wrong plan to achieve an aim" (362). In treatment of a patient with a specific disease, medical error means the wrong plan of treatment (e.g., wrong drugs are prescribed by a doctor), or failure in the implementation of the right treatment plan (e.g., wrong dispensing of drugs). Medical error can include problems in practice, products, procedures and systems (361). Medical errors include three types: one is a "close call" (namely, an error has occurred, but is found and corrected before it produces any adverse effect); another is medical error which has occurred, but does not produce an adverse effect; still another is the medical error which produces an adverse effect. Not all adverse events are to the result of medical errors. Only those which are preventable and avoidable are classified as medical errors.

There is a huge amount of medical error. In the United States, up to 98 000 people die each year as a result of medical errors; two seminal studies on medical error have shown that adverse events occur to approximately 3–4% of patients (363,364). Another study found that the

average intensive care unit patient experienced almost two errors per day (365). A 1995 study showed that problems related to the use of drugs accounted for nearly 10% of all hospital admissions, and significantly contributed to increased morbidity and mortality (366). A 1991 study of hospitals in New York State indicated that drug complications represented 19% of all adverse events, and that 45% were caused by medical errors (367). The Food and Drug Administration (FDA) receives approximately 100 000 reports per year of adverse events associated with medical devices and over 250 000 reports associated with pharmaceuticals; over one-third of these events are preventable. It is estimated that the cost associated with medical error in lost income, disability, and health care costs is as much as US$ 29 billion annually, and preventable medical errors are estimated to increase hospital costs by about US$ 2 billion in the United States (361). It is not clear what the magnitude is in other high-income countries and in low- and middle-income countries; it is to be expected that medical errors may exist in much greater magnitude in low- and middle-income countries, where they are less likely to be detected and doctors' practice is less controlled by medical standards, regulation, and law.

The report of the Institute of Medicine in the United States, entitled To err is human: building a safer health system(362) shocked the nation and attracted immediate attention by the federal government. A task force, including all relevant departments and agencies of the federal government, and a special centre (Centre for Quality Improvement and Patient Safety) was set up within the Agency of Healthcare Research and Quality. A number of research and practice efforts will be made, which include:

– defining and classifying medical errors;
– describing their frequencies and distributions;
– analysing the direct and indirect costs;
– studying the determinants;
– developing tools to reduce medical errors;
– evaluating the cost-effectiveness of these tools.

The recent switch from cost-oriented to quality-oriented health policies in the United States represents an overarching opportunity to achieve efficient performance of the system. It is believed that efforts related to the reduction of medical errors will lead to the containment of health care costs, as well as an improvement in health care quality, and thus the improvement of performance. The measures for reducing medical error being suggested or tested are: developing error reporting systems; providing non-provider-specific information on the type and magnitude of medical errors to providers, purchasers, policy-makers, and the general public; promoting the use of electronic medical order and automatic bar-code systems; and changing the look-alike and sound-alike brand names of medical products to prevent dispensing mistakes. It is still too early to provide any evidence on the effectiveness of these efforts in achieving the goals.

Monitoring quality refers to continuous, persistent and systematic oversight of the quality of health care by an external party (e.g., government bodies, and third party payers) for quality assurance. Quality of health care has long been an issue to which health care providers, purchasers and health policy-makers pay attention. Much of the attention has gone to internal quality assurance programmes, by which a health care provider is required to (or voluntarily) develops and implements a planned course of action within its organization (e.g., a hospital) to improve quality of care. Continuous monitoring of quality by an external party and linking the

results to reimbursement and non-pecuniary motivation are rare, probably because of the difficulties in measuring quality and the high cost of monitoring. Only in recent years, along with increasing concerns about quality erosion as a result of the penetration of managed care, has monitoring quality of care been put on the agenda in the United States (368,369).

Traditionally, quality assurance relies heavily on government regulations (such as licensing of medical practice, and accreditation) and qualification requirements (e.g., the existence of a quality assurance programme within the care delivery organization with specified structure, process and outcome indicators) by both public and private purchasers. These activities are, however, not carried out on a continuous, persistent and systematic basis. The satisfaction of basic requirements at the time of licensing, accreditation, and contract does not guarantee the provision of health care of good quality in the future. This provides the first reason for quality monitoring. The second reason is the increasing concern that efforts on cost control may lead to reduction in quality. In order to prevent the possible erosion of quality resulting from prospective payment and managed care (which provides incentive for the provider to contain cost, but perhaps at the sacrifice of health care quality), quality monitoring is increasingly recommended.

Quality monitoring was suggested by a RAND-UCLA research team (370) and the Joint Commission on Accreditation of Healthcare Organizations (371) when prospective payment became increasingly popular, but practical efforts did not begin until 1993, when the Health Care Financing Administration sponsored a two-year demonstration project for monitoring quality for Medicaid beneficiaries, called the Quality Assurance Reform Initiative (QARI). The demonstration involved three states (Minnesota, Ohio, and Washington) with a dual purpose: to improve the consistency of oversight of the quality of Medicaid managed care across states, and to assist states in updating and strengthening their own quality assurance systems. QARI includes two components: the promotion of internal quality assurance programmes through detailed guidelines, and external independent review of health care quality. QARI identifies 33 clinical quality indicators and 6 health service areas as concerns for the Medicaid population. This demonstration project was evaluated by Gold & Felt on its feasibility and effectiveness in improving in quality (369), but because of limitations in data and the lack of a control group, the authors did not provide information on the effectiveness of the project.

Quality of care is the heart of health provider performance; quality monitoring is a potentially effective tool for improving quality; and improvement in quality for a given budget means improvement in efficiency. Theoretically, effectiveness on allocative efficiency depends on two factors: the measurement of quality, and the use of information generated from quality monitoring. However, empirical studies of the effectiveness of quality monitoring are not to be found.

Chapter 7. Managed Care

Managed care is not an independent policy tool, because care can be managed in many ways using the tools outlined in this book. Managed care is de facto a toolkit which combines various tools to control costs and improve the quality of care. A separate section is devoted to managed care in this book, because it is difficult to separate the effect of each of the tools within managed care, and the combined effect of a number of tools is of interest. This section provides a definition of managed care and its practice in different countries, and reviews the evidence of its effects on utilization, cost, quality and health outcomes.

Definition and practice

There is no universally accepted definition for managed care (372). Some authors define managed care as the integration of financing and delivery of health services (373); others define it as care provided under the control of a third party (334).

The most comprehensive definition is that managed care refers to the mechanisms, organizational forms and techniques used by a third party (e.g government, insurer, employer or other payer) to influence the provision of health care to contain costs or improve quality of care or both, through the provision of appropriate health care services in a cost-efficient manner (373,374). In contrast to the conventional arrangement in which care is based only on patient–practitioner agreement and third party involvement is limited to paying the bill, managed care permits the third party to specify what kinds of care will be given, how and where it will be given, and how much it will cost. The third party under managed care becomes an active purchaser rather than a passive payer (334).

The mechanisms which managed care organizations (MCOs) use to control costs and improve quality, include integrating financing and provision to allow case management, and disease management through the continuum of care; transferring economic risk from insurers to providers through tools such as capitation of hospitals and physician services, bonuses and penalties for providers, and case payments; motivating providers to provide cost-effective care through payment incentives, clinical guidelines and physician profiling; restricting the autonomy of providers in medical decision-making through the use of utilization management, pre-authorization, concurrent review, retrospective review, drug formularies, generic substitution and medical practice guidelines; and limiting patient choice of providers through selective contracting with hospitals and physicians who provide high quality, cost-effective care.

The term managed care encompasses a wide variety of organizational forms. For example, the most well-known type of managed care plan is a health maintenance organization (HMO) which often has the most restrictions in choice of provider and use of services, and uses many of the tools of managed care. MCOs can be organized in various ways, such as staff models in which physicians are employees of the MCO, prepaid group practices, network models, and independent practice associations (IPA). Other more loosely managed plans include preferred provider organizations (PPOs) which offer maximum flexibility for patients and providers and use few of the tools of managed care plans; and point-of-service plans, which allow patients to select the level of restrictions they want and pay more for less restrictive use of services.

The techniques which are often used by the organizational forms described above include incentive payment (case and capitation payment for hospitals, capitation payment for physician services, withholding payment, bonus and penalties for physicians, and contracts with physicians at discounted fee schedule); utilization review (pre-authorization, concurrent review, retrospective review, and feedback); physician profiling, case management, disease management, drug formulary, generic substitution, and the use of medical practice guidelines.

Since 1973, when the United States federal government passed the Health Maintenance Organization Act to encourage the growth of managed care, managed care has developed very rapidly (373). As of 1998, as many as 75% of U.S. citizens received their health care through some form of managed care (184).

Managed care has also been adopted by other countries. Managed care was introduced to Indonesia in the middle of the 1970s, and the minister of health in 2000 strongly committed his government to expanding the managed care system through integration of financing and delivery of health care (375). Since the first prepaid group practice opened its door in Zurich in 1990, the number of HMOs has increased rapidly in Switzerland. Germany is increasingly introducing the concept of managed care in its health care system (376). HMOs have also existed in the Philippines and South Africa for many years.

Evidence

Evaluations of managed care plans have taken place exclusively in the United States. Numerous publications have provided evidence on the performance of managed care organizations, including whether managed care can reduce utilization and costs, and whether managed care can improve quality of care. However, because of the wide variety of managed care arrangements, the varying degree to which the tools in the managed care tool-kit are applied, and the complexity of the employer-funded health care system in the United States, there is no general consensus on the most effective application of managed care principles.

Utilization and cost

The majority of studies show that managed care has been able to reduce health care utilization and cost. Zwanziger & Melnick compared the trend of cost increase in the whole country with that in California, where managed care has been implemented most effectively (377). The study showed that managed care plans have been successful in inducing price competition and forcing costs down. From 1980 to 1991, the cumulative growth in real total per capita health expenditures in the United States was 63%, while in California it was only 39%. The rate of increase in real-term hospital expense per patient discharge was 4.0% in the United States and 5.6% in California over the period 1975–1982, while during the period 1982-1990 when managed care became most popular in California, the rate of increase in hospital expense per discharge increased to 4.2% nationally, but decreased to 2.4% in California.

To answer the question of whether an MCO delivers less care than the unmanaged fee-for service (FFS) system when both serve comparable populations with comparable benefits,

Manning et al. randomly assigned a group of 1580 persons to receive care free of charge from either an FFS physician of their choice or an HMO (255). This randomized controlled trial found that the rate of hospital admissions in the HMO group was 40% less than in the FFS group, although ambulatory visit rates were similar. Total health care expenditure in the HMO group was 25% less than in the FFS group and those in the HMO group received more preventive visits than the FFS group. The study concluded that the style of medical practice in HMOs was markedly less hospital-intensive and consequently, less expensive; and that the expenditures in an HMO were about 75% of those in an FFS plan (255,378).

In another randomized control trial, certain Medicaid households were assigned to remain in a traditional FFS arrangement and others were randomly selected to join a Medicaid HMO. The HMO children received equivalent numbers of check-up visits and emergency room services, but significantly fewer acute care visits (379).

Steiner & Robinson in the United Kingdom undertook a systematic literature review of the performance of MCOs based on the US literature, to draw lessons for the National Health Service (374). Their findings strongly supported the effectiveness of managed care in reducing utilization and costs. They found that:

- 11 out of 12 analyses assessing hospital admission rates reported MCO admission rates lower than, or not significantly different from FFS admission rates. Where significant differences were found, MCO rates were 26–68% lower than FFS. All but one of the 17 hospital length-of-stay analyses found shorter (one day on average) or equivalent stays for MCO patients.
- Concerning the use of discretionary care (defined as expensive tests, procedures and treatments where less expensive alternatives exist), they found 21 of 31 analyses indicated that MCOs used significantly fewer high-cost procedures than FFS plans. The case-mix adjusted ratios of MCO to FFS use of discretionary services ranged from 0.68 to 0.92 for maternity care (e.g., caesarean section), 0.23 to 0.82 for treatment of cardiovascular disease (angioplasty), and 0.65 to 0.91 for other procedures.
- In 32 of 44 observations in nine studies, covering cancer screening, immunization, child check-ups and others, significantly more patients in MCOs than in FFS received preventive care. In 10 of the 12 remaining observations, there were no observed differences, and on average, rates of preventive service were 48% higher in MCOs.
- Results were mixed in number of physician visits under managed care vs FFS. Nine studies indicated 9–50% more physician visits under managed care than FFS; 10 studies showed statistical equivalence; and four studies reported 26–68% fewer visits.
- Differences in drug prescription were not conclusive, because of inconsistent results — four observations indicated less use by MCO enrolees, three indicated more and three showed no difference.

The authors concluded that MCOs were associated with decreases in hospital admissions lengths of stay, and discretionary high-cost services, and increase in provision of preventive services. However, they were not able to draw any conclusions on differences in number of physician visits and drug prescriptions between MCOs and FFS plans.

Since 1980, only two comprehensive reviews of the managed care literature have been published in peer-reviewed journals in the United States. Both were by Miller & Luft, the first in 1994 (372) covering the period from 1980 to 1993, and the second in 1997 (380) covering the period from 1994 to 1996. Both reviews examined differences in utilization rates and costs

between MCOs and FFS plans. The reviews drew the following conclusions:

- Hospital admission rates. In the first review, of 11 observations from seven studies, HMO admission rates were lower for eight and higher for three. In the second review, the authors did not find any particular pattern in 5 studies.
- Hospital lengths of stay. In the first review of 16 observations from 13 studies, HMO length-of-stay was shorter in 15 observations compared with FFS. On balance the authors concluded that HMO plans generally had from 1% to 20% shorter hospital stay. In the second review, the authors did not find any particular pattern of differences in length of stay between MCOs and FFS.
- Hospital days per enrolee. The first review found that HMO plans consistently had lower hospital days per enrolee compared with FFS plans. In general, among the studies that showed significant differences, the number of hospital days per enrolee in HMOs was 18% to 29% less than in FFS plans.
- Use of elective costly procedures. The first review found that in 18 of 20 comparisons from nine studies, HMO plans used an average of 22% fewer procedures, tests, or treatments that were expensive or had less costly alternative interventions. In the second review, all four studies showed lower use of more costly procedures in HMOs than in FFS plans.
- Preventive care provision. In the first review, all seven studies showed that HMO enrolees consistently received more preventive tests, procedures, and examinations (such as cancer and hypertension screening tests, and breast, pelvic, rectal, and general physical examinations) or health promotion activities (such as smoking counselling practices) than FFS plan enrolees. The second review did not address the provision of preventive services.
- Physician office visits per enrolee. In the first review, half of the observations showed a higher rate of physician visits in HMOs vs FFS plans, and half showed a lower rate. In general, the evidence was insufficient to find a significant difference between the number of physician visits in HMO vs FFS plans. The second review did not find a particular pattern either.
- Total expenditure per enrolee. In the first review, the literature produced only two observations. One of them showed a 13% lower expenditure for HMO enrolees compared with FFS enrolees. In the second review, of five studies with results comparing expenditures, three showed substantially lower total expenditures for HMO enrolees (-16% to -34%) compared with FFS enrolees. Two studies showed little difference or had mixed results.

In 1998, Miller et al. focused on the effect of managed care on elderly people and reviewed the research findings of major managed care demonstration projects for the Medicare programme in the United States. They found that with few exceptions, cost savings in acute care settings were associated with managed care. Care seemed to be rationed in situations in which it could be rationed, and provided more generously when it might prove efficacious (381). According to their literature review:

- One study found that hospital inpatient use was lower for managed care enrolees compared with FFS enrolees in both admissions and days per 1000 enrolees (382).
- The total cost per month per enrolee for a capitated HMO was nearly 40% lower than for FFS (383).

- HMO enrolees had the same number of admissions, but a lower average length of hospital stay than did FFS clients, and HMOs were more likely to discharge patients to lower-cost settings (384). This study also found a 50% reduction in home health visits for HMO enrolees without any detected change in health outcome. The authors' overall conclusion was that, after adjusting for favourable selection, HMOs spent about 10.5% less for hospital, physician, home health and skilled nursing care than would have been spent for the same enrolees in the FFS system.
- HMO patients had one more physician encounter per member per year, but 26% fewer hospital days, 11% fewer admissions, and a 14% shorter average length of stay, based on a comparison of HMO Medicare patients with FFS patients from 1987 to 1989 (385).
- The annual rate of cost increase for Medicaid patients in HMOs (13.7%) was less than traditional Medicaid (15.3%) based on an analysis of costs over an 11-year period from 1983 to1994 (386).
- Average per capita expenditures in Arizona HMOs were approximately 16% lower for elderly and physically disabled long-term care Medicaid patients, compared to FFS Medicaid. Most of the savings came from reduced hospital and nursing home use (385).
- HMOs were more likely to provide elements of routine and preventive care, examine occult blood and carry out endoscopy or barium enemas among colorectal cancer patients in a study of eight Medicare demonstration projects (387).

Quality and health outcome

While the literature shows that managed care is effective in reducing utilization of care and, subsequently, costs of health care, many have expressed concerns about whether these reductions result from lower quality of care and poorer health outcomes.

In their literature review, Steiner & Robinson addressed the issues of quality of care, consumer satisfaction, and health outcomes (374):

- In terms of quality of care (measured by structure, process and outcome), they found 23 high-quality studies yielding 146 separate observations. Over two-thirds of the observations showed no significant differences between MCOs and FFS plans. Of the remaining observations, about 60% favoured treatments under managed care, and 40% favoured treatments under FFS.
- Consumer satisfaction was generally lower among managed care patients. In measurements relating to access, convenience, communications with clinicians, and perceptions of professional competence, 19 of 37 observations favoured FFS. Only six observations favoured managed care.
- With regard to health outcomes, whether defined in terms of mortality, survival time, clinical markers (e.g., blood pressure) or functional status, 84% (68 of 81) of the observations indicated no significant difference between MCOs and FFS, 11% suggested better outcomes under managed care, and only 5% indicated better outcomes in FFS plans.

In a review by Miller & Luft, 14 of 17 observations from 16 studies showed either better or equivalent quality of care for HMO enrolees compared with FFS enrolees for a wide range of

conditions, diseases, or interventions (372). The authors concluded that the HMO and FFS plans provided enrolees with roughly comparable quality of care, as measured by process or outcome measures.

In a second review by Miller & Luft, they found that five observations from four studies showed significantly better quality of care in HMOs, whereas five observations from five studies showed significantly worse HMO results (380). They concluded that HMO quality of care varied by organization and for different diseases.

Hellinger conducted another literature review on the effect of HMOs on quality of care (388). His review involved 100 studies published in peer-reviewed journals. The measures of quality included effectiveness of care, satisfaction with care, and access to care. He concluded that managed care did not decrease the overall effectiveness of care. Evidence suggested, however, that managed care might adversely affect the health of some vulnerable subpopulations (e.g., the elderly and the chronically ill). Evidence also suggested that enrolees in managed care plans were less satisfied with their care and had more problems accessing specialized services.

Focusing on the elderly population, Miller et al. reviewed the research on the association between managed care and quality of care for Medicare patients (381). According to their review:

- Medicare HMO beneficiaries were less satisfied with the care process, plan access, provider choice, and perceived quality and health outcomes than FFS beneficiaries. However, they were significantly more satisfied with their costs(384).
- HMO enrolees were more likely to obtain cancer screening services, and were diagnosed at earlier stages for cancer sites, for which effective screening services were available (389).
- Comparisons of health outcomes of acute care between HMOs and FFS yielded mixed results. Some acute care outcomes were no different in HMOs compared with FFS (390-393), and some acute outcomes in HMOs were worse (384,392-394).
- The authors also noted previous studies that suggest that managed care did worse on a variety of process and outcome indictors in treating conditions of a chronically ill, frail elderly population in both nursing homes and home health settings.

Rogers, Wells & Meredith focused on managed care quality for the mentally ill. They found that managed care mental health carve-outs (plans in which mental illnesses are managed separately from other acute illnesses) may not be warranted and stated, "Mental health advocates and clinicians have long maintained that managed mental health programs achieve cost savings, at least in part, by limiting benefits and making it difficult for patients to access care" and "Although the research findings are mixed with regard to capitation for mental health services, the RAND Medical Outcome Study found that patients treated for depression in Independent Practice Association-model HMOs experienced significantly worse outcomes" (395).

Summary

The principles underlying managed care suggest that it should reduce hospital use and costs, emphasize primary and preventive care, discourage inappropriate use of services, and substitute less expensive alternatives such as ambulatory, home health, and nursing home care for care in

more expensive delivery settings. In the long-term care setting, managed care should be expected to prevent hospitalization and institutionalization of patients by supporting them in assisted living and home care settings, improving coordination of care, and providing patients with services uniquely required by a geriatric population (381).

Empirical evidence on utilization and costs supports these claims. The absolute majority of studies showed that managed care can reduce utilization and costs, through a better mix of cost-effective services and fewer costly discretionary interventions. Some studies challenge these findings and contend that the differences in utilization and costs between managed care and FFS plans are a result of differences in the case-mix of patients by managed care plans (i.e.healthier people tend to select MCOs). For example, a comparative study of 308 HMO patients with 529 FFS home health patients found that the case-mix of Medicare FFS patients compared with Medicare HMO patients included more impairments in activities of daily living (ADL) and various physiological conditions (396). Despite these criticisms, however, the consistent results among randomized controlled trials and case-mix adjusted analyses show that there is little doubt about the validity of the conclusions related to utilization, intervention mix, and cost.

What is still controversial is whether the implementation of managed care can improve health outcomes. The evidence cannot support a firm conclusion that managed care effectively improves the quality of care. The major reason for this may be that managed care plans in the United States have put too much emphasis on cost containment and too little on quality improvement. In addition, as suggested by Miller & Luft, quality improvement might be associated with particular types of managed care organizations and particular diseases, and quality assurance capacities vary significantly among individual MCOs. Also, an association between resource use and health outcome for many diseases is often tenuous, (e.g., the amount of resource input may not be associated with improvements in health outcomes for advanced liver cancer patients). Further, the lack of a uniform measure of quality makes comparisons difficult (380).

While the proponents of managed care tend to search for favourable evidence (397), opponents provide strong criticism of managed care. These criticisms include unnecessary constraints over the choice of providers, potential harm to the doctor–patient relationship, reduction in access to specialists and diagnostic tests, and possible harm to quality of care and health outcomes (398). Evidence shows, however, that although managed care is not as good as some would like to believe, it is better than the criticism levelled at it. There is no firm evidence that managed care results in less access to necessary care, and there is no consensus on the negative effects of managed care on quality of care and health outcomes.

In conclusion, managed care is the term for a toolkit which encompasses numerous organizational forms, mechanisms and techniques. Despite literally thousands of publications since 1990 whose subject is some component of the managed care approach, it is still not possible to answer fundamental questions about the independent contribution of each type of managed care plan and each tool to the general performance of managed care plans (374). This is a ripe area for future research.

Chapter 8. Summary of Policy Tools

The effectiveness of policy tools has been reviewed against their specified objectives, which include one or several of the following: to correct maldistribution of providers; to contain the cost of health care; to improve quality of care; to increase access to basic interventions; to reduce the utilization of inappropriate care; and to restrict availability of cost-ineffective interventions.

Although none of the policy tools have as an explicit objective to improve the allocative efficiency of health interventions, which is the interest of this study, all of them may have secondary effects on allocative efficiency. For example, an increase in access is linked to allocation of health interventions; correcting a maldistribution of providers is associated with improvement in the location of providers and the types of services provided; cost containment may be a result of reduction in use of inappropriate care and improvement of intervention mix; the increase in availability of basic interventions is directly related to improvement of the mix of health interventions; the reduction in utilization of inappropriate care is in fact improvement of intervention mix; and the restriction of availability of cost-ineffective interventions means improvement of allocative intervention.

Tool summary

Table 1 provides a summary review of possible policy tools. Their effectiveness is divided into four dimensions: cost containment, quality and health outcome improvement, changes in utilization, and improvement in intervention mix (allocative efficiency). Where there is empirical evidence, it is specified. If there is no evidence, theoretical effectiveness is summarized. The evidence is divided into two types: positive evidence, defined here as evidence of effectiveness in achieving what is intended; and negative evidence of effects opposite to what is intended.

Table 1 Summary of the effectiveness of policy tools on allocative efficiency of health interventions

POLICY TOOLS AND OBJECTIVES	COSTS	QUALITY AND HEALTH OUTCOME
1. Budgeting input for public provision: to increase access to care	No evidence; theoretically, where government provision dominates the system, government can control costs	No evidence; theoretically, depending on adequacy of budget and its allocation amongst alternative uses
2. Strengthening medical education: to increase access to care and to correct manpower imbalance	Not applicable	No evidence; theoretically, quality of care and health outcome may improve through improvement of access to care
3. Strengthening health facilities: to increase access to care and to correct provider imbalance	Not applicable	No evidence; theoretically, quality of care and health outcome may improve through improvement of access to care
4. Training grant incentives: to correct misdistribution and improve access to care	Not applicable	No evidence; theoretically, quality of care and health outcome may improve through improvement of access to care
5. Funding cost-effective products: to increase access to these products	Not applicable	No evidence; theoretically, may improve through increase in access to cost-effective products
6. Certificate-of-Need: to control capital investment and cost of care	Negative evidence; in theory, control of capital may lead to reduction in costs	No evidence; theoretically, perceived quality may suffer; but may not necessarily lead to worse health outcome
7. Capital cap: to control capital investment and cost of care	Positive evidence; there is evidence capital cap + Certificate-of-Need is effective in cost control	No evidence; theoretically, perceived quality may suffer; but may not necessarily lead to worse health outcome
8. Capital moratorium: to control cost of care	Negative evidence: temporarily effective, but not effective in the long run	No evidence
9. Technology regulation: to restrict availability of cost-ineffective interventions and to control the cost of care	No evidence; theoretically, it can control costs through providing information and translating it into policy	No evidence; theoretically, they may improve through control of dubious technology
10.Fee-for-service: traditional method, increasingly used to encourage provision of certain services	Negative evidence: increase in health care costs because of the incentive to provide more and costlier health care	No evidence; theoretically, increase in perceived quality, but health outcome depends on the extent of overprovision
11. Capitation: to control cost	Positive evidence: successful in cost containment through removal of incentive to provide more and costlier care	No evidence; theoretically under-provision may reduce quality and health outcome
12. Salary: traditional payment method without incentive to overprovide or underprovide	No evidence; theoretically, has neutral incentive for provision, and lower cost than FFS, but higher cost than capitation	No evidence; theoretically under-provision may reduce quality and health outcome
13. Daily payment: to control cost	No evidence; theoretically, reduction in cost/day, but not necessarily cost/patient	No evidence
14. Case payment: to control cost	Controversial: reduction in inpatient cost, not overall cost because of cost shifting	Controversial; theoretically, earlier discharge of patients (the problem of "quicker and sicker release") and reduced necessary and unnecessary provision

UTILIZATION	INTERVENTION MIX
No evidence; in theory, pattern of use depends on provision; it also depends on presence and absence of charges	Negative evidence: in theory, policy-makers are in a position to optimize intervention mix, but it is rarely the case
No evidence; theoretically, utilization will increase because of increase in supply and access	Negative evidence in the US; theoretically, balanced manpower may lead to balanced intervention mix
No evidence; theoretically, utilization will increase through increase in supply and access	Negative evidence in the US; theoretically, balanced facility may lead to balanced intervention mix
No evidence; theoretically, utilization in underserved areas will increase because of increase in supply and access	No evidence; theoretically, intervention mix may improve through improvement in the mix of providers
No evidence; theoretically, utilization of cost-effective products will increase	No evidence; theoretically, intervention mix may improve because of improved availability of cost-effective products
No evidence; theoretically, utilization of capital-related services may fall if capital investment is controlled	No evidence; theoretically, there will be improvement in interventions mix if capital investment is controlled
No evidence; theoretically, utilization of capital-related services may decrease if capital is controlled	No evidence; theoretically, control of inappropriate capital investment may improve intervention mix
No evidence	No evidence
No evidence; theoretically, if successful, technology regulation can reduce use of cost-ineffective interventions	No evidence; theoretically, if inappropriate technologies are controlled, there will be an improved intervention mix
Negative evidence; overprovision of health interventions, especially for profitable services	No evidence; theoretically, depending on the relative level of fees different interventions in relation to their costs
Positive evidence: reduction in more costly interventions and increase in less costly interventions	Changes in intervention mix, but it is unclear whether these mean improvement in allocative efficiency
Positive evidence; salary payment is associated with lower use of tests, referrals and procedures	Changes in intervention mix but it is unclear whether these mean improvement in allocative efficiency
Negative evidence: associated with long length of stay and large number of admissions	No evidence; theoretically, less care per hospital day, but more days of stay and admissions
Controversial: reduction in length of stay, and increase in the number of admissions; and shifting to outpatient settings	No evidence; theoretically, less cost-ineffective intervention for inpatients, and more outpatient services

Table 1 Summary of the effectiveness of policy tools on allocative efficiency of health interventions (continued)

POLICY TOOLS AND OBJECTIVES	COSTS	QUALITY AND HEALTH OUTCOME
15. Line budget: to control allocation of budget	No evidence	No evidence; theoretically, possible rigidity in provision
16. Global budget: to control cost	Positive evidence: a powerful tool to control cost	No evidence; theoretically, tight control over cost may reduce quality of care and may have negative effect on health
17. Performance-related pay: to improve performance	No evidence; theoretically, depending on whether cost containment is included in the performance measure	No evidence; theoretically, if performance is measured by quality and health outcome, there will be improvement
18. Rate setting: to control cost and increase access to care	Positive evidence: reduction in price can control cost despite increase in volume	No evidence; theoretically, low perceived quality if price is too low
19. Fee structuring: to control cost and encourage the provision of certain services	Positive evidence: can control both price and volume, thus the containment of cost	No evidence
20. Separation of dispensing from prescribing: to control cost fo drugs	Some evidence: dispensing doctors prescribe more and costlier drugs than non-dispensing doctors, and this drives drug cost high	No sound evidence; theoretically, neutral incentive to prescribe will improve quality and health outcome
21. Payer integration: to control cost	Positive evidence: an effective tool for controlling national health care expenditure	No evidence; theoretically quality and health may reduce unless mix of use improves
22. User fees: to generate more revenue, improve quality and increase access	Not applicable	Controversial; theoretically, depending on whether revenue from charges is sufficient and retained
23. Cost sharing: to control cost	Controversial, but likely to have limited effect on cost control	Negative evidence: negative effect on health, particularly for the poor
24. Rationing: to control cost by limiting cost-ineffective care and increase access to cost-effective care	Negative evidence: failure in cost control in Oregon	No evidence; theoretically, possible to improve health through improvement in access to basic care by more people
25. Essential package of interventions: increase availability and access to basic interventions	Not applicable	No evidence; theoretically possible to improve quality and health through increase in access to basic care by more people
26. Essential drug list: to increase availability of essential drugs and to control cost	No evidence; theoretically the increase in rational use of drugs will lead to decrease in costs	No evidence; theoretically, can improve quality and health through increase in access to essential drugs
27. Medical practice guideline: to control cost and improve quality by limiting the use of inappropriate care	Positive evidence: reduction in health care cost	Positive evidence: improvement in quality of care and health outcome
28. Prescription formulary: to control cost through the reduction in inappropriate prescription of costly drugs.	Controversial evidence: reduce cost for drugs, but not necessarily overall cost for care	No evidence; theoretically, possible through reduction in unnecessary drugs
29. Utilization review: to control cost and improve quality through reducing inappropriate care	Controversial evidence: reduction in cost of care, but increase in cost of utilization review	No evidence; theoretically, possible through decrease in inappropriate interventions
30. Managed care: to control cost through intervention of providers' provision decisions.	Positive evidence: reduction in overall cost for health care	Controversial: without negative effect on health (except for vulnerable population), but less consumer satisfaction

UTILIZATION	INTERVENTION MIX
No evidence	No evidence; theoretically, provision may not reflect the need and cost-effectiveness
Negative evidence: overuse of outpatient services; and less access to hospital care when running out of budgeted funds	No evidence; theoretically, possible increase in cost-effective services and decrease in cost-ineffective ones
Negative evidence: quantity- related measure led to more provision, and revenue-related measure led to overprovision	No evidence; theoretically, if performance is measured by health/cost, intervention mix will be improved
No evidence; theoretically, depending on who pays for care	No evidence; theoretically, the mix of interventions depends on fee structure
Positive evidence: increase in use of primary care and decrease in use of elective procedures	Changes in interventions mix observed; but it is unclear whether these mean improvement in allocative efficiency
Some evidence: dispensing doctors prescribe more and costlier drugs than non-dispensing doctors	No evidence; theoretically, the overall cost-effectiveness of drug use is expected to increase
No evidence: theoretically, increase in use of basic care, and decrease in elective care	No evidence; theoretically, fewer elective services under single-payer systems
Negative evidence: reduction in utilization of all services, especially for the poor	No evidence; theoretically, depending on fee structure and implementation of exemption for the poor
Negative evidence: reduction in total utilization of both necessary and unnecessary care	Negative evidence: reduction of both cost-effective and cost-ineffective interventions
No evidence; theoretically, possible to reduce use of cost-ineffective services	No evidence; theoretically, possible to improve the mix of interventions through increase in coverage of essential care
No evidence; theoretically, possible to increase the use of basic care	No evidence; theoretically, possible to improve the mix of interventions through increase in access to essential care
Positive evidence: increase in use of essential drugs and decrease in use of non-essential drugs	Changes in the mix of pharmaceutical drugs: more essential drugs and less non-essential drugs
Positive evidence: reduction in use of cost-ineffective interventions	Positive evidence: improved intervention mix through increase in cost-effective care
Controversial evidence: reduction in use of drugs, but not known whether necessary or unnecessary drugs	No evidence: theoretically, possible if there is a decrease in unnecessary drugs, and no decrease in necessary ones
Positive evidence: reduction of hospital use, and the use of expensive drugs	Changes in intervention mix through reduction in elective care
Positive evidence: reduction in utilization of elective services and increase in preventive care	Changes in intervention mix reduction in costly and selective interventions

Observations

The above summary of the policy tools can be elaborated by several observations. First, although there is some evidence related to the description of the problems that the policies were meant to deal with, none of the policies were based on sound evidence (from a pilot study or review of the effects of alternative policy tools) to evaluate their effectiveness and specify the conditions for success. Rather, these policies were based on theoretical predictions.

Second, since the objective of most policy tools is to control costs of health care, and most studies evaluate a particular policy against its goal of cost containment, evidence on impacts is not comprehensive. Most evaluations focus on cost and utilization, and rarely take into consideration quality, health outcome, and efficiency. The studies which focus only on cost containment provide little information on whether the policy tools can be used to improve the overall performance of the health care system.

Third, the effectiveness of most policy tools is not as theoretically predicted. The reasons may be multiple: the principle is good, but the policy is proved not to be feasible for political reasons (e.g., rationing, and budgeting inputs for public provision); or the implementation lacks continuous evaluation and information feedback regarding the problem the policy purports to solve. As a result, the policy is still in place while the problem has changed, and continuation of the policy creates new problems (e.g., strengthening medical education programmes, and the Hill-Burton project in the United States). The design and implementation may depart from the objectives through neglect of the conditions for success. For example, Certificate-of-Need is effective only with a capital cap; diagnosis related grouping is effective only if it is an inclusive payment under an all-payer system; and the effectiveness of performance-related pay depends on how performance is measured.

Fourth, several policy tools (capital cap, global budget, rate setting, and payer integration) are powerful tools for controlling cost, but there is no evidence whether these tools can improve allocative efficiency, although theoretically it is possible. There might be three reasons: there is an effect, but evaluation has yet to find it; the potential to improve allocative efficiency has been ignored; or the complementary policies to improve efficiency are not in place.

Fifth, except for medical practice guidelines, evidence of effects on quality and health outcome is hardly available. One possible reason is that the policy objective is usually not to improve the quality of care, and studies addressing the policy's effect on quality and health outcome are not of interest. Another reason is technical difficulty in the measurement of quality. This is not only because of the difficulty in developing a valid measure, but also because the lack of a uniform measure of quality makes different regimens hardly comparable. In addition, health is determined by many factors, and it is difficult to set up a causal link between health outcome and the implementation of a policy. Another reason is that there is no observed change in quality and health outcome. The importance of lack of change is often neglected because it is mistakenly believed that this means there is no positive effect on system performance. If sound evidence shows that the implementation of a policy leads to no change in quality and health outcome, but can control the overall cost of health care, the policy tool is surely of great policy significance, because it means efficiency improvement — more health gain from a given amount spent.

Sixth, although none of the policy tools is purposefully designed to improve allocative efficiency, some of them have a great potential to do so. These policies include capitation payment, salary payment, fee structuring, essential drug lists, medical practice guidelines, utilization review, and managed care. Research is needed to develop the tools with potential into means of efficiency improvement. Studies on how to combine the various tools into a powerful toolkit to improve the efficiency performance of the health care system are also needed.

Combinations of policy tools

The lack of information on the cost-effectiveness of interventions and the absence of uniform quality and health outcome measure across different interventions constitute major technical barriers both for designing a policy tool and for finding evidence of the effect of the tool on allocative efficiency. Information generation and methodological development are further challenges that have to be tackled.

All the policy tools reviewed can be categorized into three groups: tools related to the control and distribution of inputs for the production of interventions; tools related to the management of provision of health interventions; and tools related to purchasing health interventions.

The control of intervention inputs includes all the policy measures that specify the type, quantity, and location of various inputs for the production of interventions. The rationale for this control is that the type, quantity, and location of inputs may determine the type, quantity, and location of intervention output, and positive change in the former may lead to improvement in the latter. Related tools include technology regulation, strengthening medical education, Certificate-of-Need, capital cap, capital moratorium, strengthening health facilities, training grant incentives, and budgeting inputs for public provision. It is expected that the implementation of these policy tools can result in the improvement of the second and the third steps of resources allocation, which then has a positive effect on the later steps.

Managing intervention provision includes all the policy measures that specify the type or quantity of health interventions that should be provided. The rationale is that there exist over-provision of cost-ineffective interventions and underprovision of cost-effective health interventions; and health care providers (particularly private providers and public providers with financial autonomy) are not likely to provide the best mix without policy intervention. The tools related to managing intervention provision include funding cost-effective products, essential packages of interventions, essential drug lists, medical practice guidelines, prescription formularies, and utilization review. It is expected that specifying what should be provided and what should not, with penalties and rewards, can improve the overall efficiency of the health care system.

Purchasing of interventions reacts to all the policy measures that affect the types and quantities of health interventions utilized or provided. In this purchasing process three parties are involved: the third party payers, the users, and the providers. In addition, policy-makers usually are intensively involved in this process to set the rules for purchasing care and strengthening the efficient performance of health care systems. Policy tools related to purchasing can be divided into three groups: those that influence providers, namely providing incentives for providers to delivery socially desirable health interventions to the users; those that affect the users, including economic incentives and information for users; and those that determine the

packages of services that are purchased by third party payers. The rationale for designing policy tools to affect purchasing of interventions is that purchasers, providers and users will not make cost-effective choices unless they are motivated to do so. In addition, information for users and purchasers as well as providers will strengthen the purchasing process. The policy tools related to purchasing interventions are mostly designed by third party purchasers (private or government insurers) who have been under a great deal of pressure to contain costs. These tools include user fees, cost sharing, rationing, fee-for-service, capitation, salary, daily payment, case payment, line budgets, global budgets, profit-related pay, rate-setting, fee structuring, separating dispensing from prescribing, payer integration, and managed care. It is expected that strategic purchasing using these tools can reduce the provision of cost-ineffective health interventions, and the reduced cost for inappropriate care can be shifted to the provision of cost-effective interventions.

Targeting limited government funds to the provision of cost-effective health interventions to deal with the diseases and risk factors that bring about the greatest health burden can improve the allocative efficiency of the government budget. Internationally, there is anecdotal evidence that government-sponsored special programmes to cope with those diseases are successful in transferring resources to more cost-effective alternatives. However, it seems that a systematic way of doing this has not been fully developed. This is hindered not only by a lack of technical support to provide scientific evidence for prioritizing health conditions and interventions, but also by a lack of willingness, ability and methods for the government to realize the real transfer of resources. To spend government money wisely and direct government resources to the health conditions and interventions that yield the highest return, two questions have to be answered by both researchers and policy-makers.

The first question is: What are the health conditions and risk factors that currently and in the future cause the greatest burden to the population? Research is needed to rank the diseases and risk factors according to their size of burden — one of such approaches is disability adjusted life years (DALYs) lost to diseases and injuries. Methods for prioritizing diseases, injuries and risk factors have been provided, and are obtaining increasing attention from both researchers and policy-makers (12,399,400). However, because of the lack of data, or lack of well-trained researchers, or lack of government attention, research related to prioritizing health conditions at country level is not widespread. The important point is that only if diseases and injuries are prioritized, can it be possible to prioritize the allocation of a limited government budget.

The second question is: What are the most cost-effective interventions for the improvement of population health? To answer this question, a list of health interventions (including services, procedures and medical products) and their cost-effectiveness ratios is needed. However, cost-effectiveness analysis to evaluate health interventions is dominated by studies of prospective new interventions compared with current practice (400); and the effectiveness measures range from physiological and biological measures (blood pressure, the size of tumour, the level of blood sugar, etc.) to more aggregated measures using generic and specific indicators of the quality of life. These are not consistent across different types of interventions, which does not allow for comparison of cost-effectiveness across all possible interventions. The key point is that only if the effectiveness measure is standardized and cost-effectiveness ratios are available for all or at least the major interventions, can the priority interventions be determined.

As an international effort, WHO has been seeking to provide evidence-based information to decision-makers for setting priorities and improving the performance of their health systems. The Global Programme on Evidence for Health Policy (GPE) of WHO is assembling regional databases on the costs, impact on population health and cost-effectiveness of key health interventions. This work, known as WHO-CHOICE, started in 1998 with the development of standard tools and methods and is now in the phase of collecting and analysing the necessary data on costs and outcomes.[4] The objectives of WHO-CHOICE are to:

- develop a standardized method for cost-effectiveness analysis that can be applied to all interventions in different settings;
- develop and disseminate tools required to assess intervention costs and impacts at the population level;
- determine the costs and effectiveness of a wide range of health interventions, presented with probabilistic uncertainty analysis;
- summarize the results in regional databases that will be available on the Internet;
- assist policy-makers and other stakeholders to interpret and use the evidence.

Policy-makers are concerned with two questions requiring evidence on costs and effects. First: Do the resources currently devoted to health achieve as much as they could? To answer this question, the costs and effects of all interventions currently employed must be compared with the costs and effects of alternatives. Reallocating resources from inefficient to efficient interventions can increase population health with no change in costs. Second: How best to use additional resources if they become available? This type of analysis is critical for ensuring that as societies become wealthier, additional resources are well used. But it is pointless asking this type of question if the current mix of interventions is inefficient — both questions need to be asked together.

WHO-CHOICE permits both questions to be asked and both types of analyses to be undertaken simultaneously. The generalized cost-effectiveness analysis forms the basis of the WHO-CHOICE approach. This method allows existing and new interventions to be analysed at the same time. Previous cost-effectiveness analyses have been restricted to assessing the efficiency of adding a single new intervention to the existing set, or replacing one existing intervention with an alternative. Using WHO-CHOICE, theanalyst is no longer constrained by what is already being done, and policy-makers can revisit and revise past choices if necessary and feasible. They will have a rational basis for deciding to reallocate resources between interventions to achieve social objectives.

WHO-CHOICE allows comparison of current interventions together with interventions being contemplated for implementation. It takes into account, from the health systems perspective, synergies between interventions on costs and effectiveness.

Prioritization of health conditions and cost-effectiveness of health interventions in a comparable way are both necessary, but not sufficient. In the search for cost-effective interventions, two points have to be made clear. One is that priority health conditions may not be a priority of resource allocation if there are no cost-effective interventions for them. Another is that the same intervention can be used for different health conditions; and if it is cost-effective for one disease, it may not be cost-effective for another. The answer to either question can capture only part of the priority picture. One approach is to link the list of health conditions with the list of health interventions, and work out condition–intervention pairs. The

[4] Web site: http://www3.who.int/whosis/menu.cfm?path=evidence,cea&language=english

linked information can be put in the form of a matrix and provided to policy-makers for resource allocation decisions. Table 2 shows a simplified form of such a matrix, in which:

- all diseases are categorized into three groups (high, medium, and low priority diseases);
- all interventions are categorized into three groups (highly cost-effective, moderately cost-effective, and cost-ineffective interventions);
- the matrix which links the two dimensions consists of nine cells which can provide useful information for decision-makers to allocate limited government budgets.

Table 2 indicates that the government budget can be strongly recommended for high priority diseases with cost-effective interventions; the government budget should not be spent for low priority diseases with only cost-ineffective interventions; and on other occasions decision-makers have to decide between using government funds to fight priority diseases regardless of the cost-effectiveness of available interventions, or for cost-effective interventions regardless of the priority of diseases. Our suggestion is shown in the table by the number of "+" and of "-", the sum of both being 5. The higher the number of "+", the more strongly we recommend using the government budget; and the higher the number of "-", the more strongly we recommend against it. The final decision will, however, be based on the preference of the policy-makers, within a specific economic and political context. Research to develop such a matrix further (or an even more complex one) is strongly recommended.

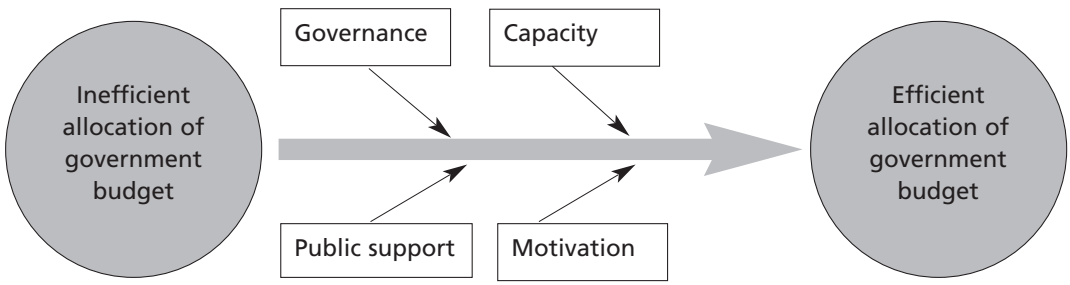
Table 2. Disease–interventions matrix: linking prioritization of health conditions with prioritization of health interventions in resource allocation decisions

	High priority diseases	Medium priority diseases	Low priority diseases
Cost-effective interventions	+ + + + +	+ + + +	+ + +
		-	- -
Moderately cost-effective interventions	+ + + +	+ + +	+ +
	-	- -	- - -
Cost-ineffective interventions	+ + +	+	
	- - -	- - - -	- - - - -

Note: + recommended use of government budget; - recommendation against government budget use.

The next question is how to target the limited government budget to priority conditions and interventions, as suggested in Table 2. People may think that once scientific information is provided on the priority use of government funds, the actual allocation should be easy. In fact, this is the most difficult step of government budget allocation. It is evident that most health authorities in both developing and high-income countries can point out (not necessarily scientifically) several priority conditions (e.g., AIDS, tropical diseases, injuries, and cancer) and categories of cost-effective interventions (e.g., primary care and preventive services), but the government budget is still difficult to divert from tertiary care to these cost-effective interventions to deal with the priority diseases. Here, several requirements for a better government budget allocation can be suggested, which include promoting good governance, strengthening capacity, mobilizing the support of the public, and providing motivation (see Figure 6).

Figure 6. The requirements for promoting the allocative efficiency of government health budgets



Stewardship is of key importance. The World Health Report 2000 defines it as careful and responsible management of the well-being of the population, and is a hierarchically higher concept than governance and administration (1). However, sub-optimal decisions are often made by governments because of political considerations. These may have negative consequences for health system performance.

Lack of capacity is a barrier to the government making rational resource allocation decisions, especially in low- and middle-income countries. Capacity, as here defined, is the ability of key decision-makers to overcome the technical constraints that limit the likelihood of making better decisions. Policy-makers should be able to digest all the information and evidence available in making their decisions; conduct effective dialogue with other stakeholders towards better decisions; make clear what is the gap between the best possible (or desired) allocation and the actual budget. This should be a reiterative process.

An informed public can use its collective power to influence government decisions. Transparency in the budgeting process and a clear tracing of resource flows either through public expenditure reviews or national health accounts could show where funds actually go and who actually pays for the services. An informed public serves as a check and balance against political pressures.

Another requirement is motivation for the government to allocate its health budget efficiently. This is particularly important in decentralized health care systems, where the central government serves as a steward, and the local governments control the allocation of budgets. In such a scenario, different mechanisms are needed to motivate local governments to use their health budgets more efficiently. For example, the central government can provide a special budget for provision of cost-effective interventions, but require matching funds from the local governments. The central government can set a standard for budget allocation, and reward

local governments for meeting this standard. These mechanisms may be particularly important in a decentralized context.

Once the decision on overall allocation of the government budget amongst various diseases and interventions is decided, policy-makers have to decide how the allocated budget is used. It can provide defined packages of health interventions, or special disease control programmes. In either case, policy-makers have to take either an input-oriented or an output-oriented approach in funding the defined interventions. They also have to decide to use public or private provision systems, or both. The analysis here strongly suggests an output-oriented approach, by which government provides funds for providers to deliver specified types and amounts of health interventions for the target population. Input-oriented funding, by which government provides funds for inputs (facilities, equipment, and staff) without specification of the output, is likely to result in productive inefficiency and the provision of health interventions that depart from priority diseases and cost-effective interventions. It is very difficult to recommend whether to use a public provision system or the private sector to provide priority health interventions to deal with priority diseases, because the decision largely depends on country context. The principle is that government should use its limited health budget for the provision of the maximum amount of defined health interventions for its target population, rather than protecting the existing public providers regardless of their productive efficiency. Public provision exists in many countries, but it is problematic because of the productive efficiency problem worldwide and abuse of the system by the top elite in low- and middle-income countries. If a government budget is output-oriented, and government contract services are based on bidding among various providers, this not only can promote allocative efficiency, but also force providers to improve their productive efficiency.

Beginning in the early 1980s, there has been a decreasing trend in direct public provision, an increase in privatization, and an increase in insurance-oriented government funding. Although the government budget for the direct provision of health care is still important, it represents a decreasing share of the total health spending of a society. In consequence, much attention is focused on developing mechanisms to regulate providers so they will deliver a socially desirable mix of health interventions.

The increased availability of essential drugs can increase the use of such drugs. By the same token, improvement in availability of essential services, such as immunizations and prenatal care, can increase their utilization. China, Sri Lanka, and some African countries were successful in increasing the use of cost-effective health interventions by increasing their availability through government-funded public provision systems. But when charges which covered the cost of health interventions were imposed, there was a decrease in use by the people who were not able to pay. Apparently availability of essential health interventions is necessary, but not sufficient for their actual utilization. Additional measures need to be taken for financing these health interventions for those who are not able to pay for them.

The practical problems of how to increase availability remain. The alternatives include government provision of cost-effective interventions, making them available to everyone; a highly regulated private medical market in which the location of providers is controlled (e.g., a quota for the number of physicians per 1000 population in a small area, after which no more are allowed to enter the market), and the drugs and services provided by particular providers are specified and regulated; economic motivation for health workers to practise in underserved areas; and economic motivation by reimbursing cost-effective services more generously or by

not reimbursing for cost-ineffective interventions. All these options exist in reality, but there is little literature on the experiences of various countries and on the effectiveness of these alternatives. Research needs to be done to provide evidence on how to improve availability of cost-effective health interventions and on the selection of cost-effective strategies.

Health care providers are major decision-makers of what health interventions are provided and to what kind of people. This simple fact had not been stressed by policy-makers and researchers, until cost escalation and low economic efficiency of health resources use became apparent in the early 1970s. Nowadays, clinical governance, medical guidelines, pathways, and payment systems have become buzzwords in health policy and systems research and in health policy-making. These tools for intervening in medical decision-making are based on two facts. One is that providers often recommend treatment for individual patients which is costly and only marginally effective, without envisaging how resources should be allocated among health interventions to achieve the highest possible level of health. This is particularly the case when patients face a zero or very low price at the point of use and providers are reimbursed for whatever services are provided. Another fact is that doctors are economic beings, and income is an important argument in their utility function. When there is a link between provision and income, doctors are likely to provide more and costlier health interventions when fewer and less costly interventions with the same effect are available. Thus the importance of providers' behaviour is attracting increasing attention of health policy-makers and researchers, who believe that providers will not allocate health interventions in a socially desirable manner unless their decisions are influenced by outside forces, among which are participating in medical decisions and motivating health care providers.

All types of purchasers can participate in medical decisions, including consumers, third party purchasers, and governments. The ways of participation differ depending on the participants, but all of them can involve themselves in medical decision-making. The degree of participation of consumers depends to a large extent on the stock of information the consumers already have and the provision of information by providers at the point of services.

Information for consumers can be divided into four types, each of which plays a different role in the health care market. The first is information for consumers to help prevent disease and manage their own health. The objective of providing this type of information is to promote population health and reduce the requirement for health resources. The second type of information is related to strengthening the power of patients in choosing providers. The objective here is to empower patients' choice of providers, and promote competition among providers based on quality and price. The third type of information is medical knowledge that enables patients' participation in medical decision-making in their interactions with providers. The objective is to have practitioners' medical recommendations reflect the preference of patients and to increase patient satisfaction. The fourth type is the information which can be used by consumers in choosing their health insurance plans. The objective of providing this information is to promote competition among insurers. Although all four types of information are relevant to allocative efficiency, only the second and third types fall within the scope of this work. In the case of an integrated purchasing model, discussion will extend to informing consumers about health plans which are responsible for both financing and provision.

Consumers' lack of information is one of the major reasons why the health care market is different from the markets for common goods and services which are driven by the demand of consumers, and where consumers have enough information to value the goods and services

they consume and to make rational choices to reach the highest level utility for the money spent. The lack of information is manifested in the following aspects.

- The prices of health interventions (including services, drugs and other medical products) are often unknown by the consumers until after the treatment. Even when patients are told the prices by a provider before health interventions are delivered, patients generally do not know the prices of the same health interventions offered by other providers, which is a necessary condition for patients' demand to work efficiently in the market.
- The quality of care is often unknown by the consumers. The quality consumers perceive (amenity, seniority of doctors, the attitude of providers), may not reflect the technical quality which is related to health outcomes. Patients often classify quality based on the level of administration and the scale of providers: thus patients are likely to care at higher-level hospitals when it is available at lower-level hospitals at lower direct and indirect costs.
- The health interventions that should be used are largely unknown to the consumers. On most occasions, patients have to depend totally on doctors' recommendations, which may be biased because of the limitation of doctors' knowledge of the cost-effectiveness of alternative interventions, and because of the possible economic motivation for the doctors to provide either more or less than medically needed care depending on the way providers are paid.
- The benefits of interventions — their effect on health outcomes — are not known by the consumers. The effectiveness of a particular intervention is influenced by many factors (random events, patients' capacity of recovery, and specific condition of diseases) that are beyond the control of both providers and consumers (401). Even doctors cannot predict the definite outcome of an intervention. A second reason is that patients are usually not given the information that is known by providers. The lack of information about benefits constitutes a major barrier for consumers' rational choice.

Ideally, if consumers were perfectly informed, the health care system would not have so much trouble as it has today. Unfortunately, no one can expect that information for consumers to purchase health care can be enhanced to a level comparable to that for purchasing apples in the market. This is by no means to say efforts to inform consumers are futile. Theoretically, better-informed consumers are in a better position to choose interventions that are cost-effective. Efforts to inform consumers can be made by government through media, by providers and by third party purchasers actively or passively (with regulatory requirement).

Partly because of the difficulty of informing consumers, and partly because of the lack of evidence that participation of better-informed consumers can lead to better medical decisions, participation by collective purchasers (either third parties or government agencies) has become much more widespread. Medical decisions regarding what types of health interventions to provide to what types of people are traditionally the results of interactions between providers and patients. The collective purchasers' function is traditionally nothing more than paying the bills. Because patients are in a weak position in making medical decisions, medical practitioners traditionally enjoy a great deal of autonomy. Driven by the escalation of heath care costs and in recognition of the likelihood of overprovision of health care, collective purchasers are increasingly involved in medical decision-making.

In contrast to individual consumers, collective purchasers are in a better position to participate in medical decisions and influence the choice of health interventions, because they are large

purchasers and have the economic power to influence providers' decisions, and because they can employ expert opinion and knowledge to decide what to purchase.

Participating in medical decisions

There is an obvious overlap between the tools related to managing intervention provision and those related to participating in medical decisions. Medical practice guidelines, prescription formularies, and feedback have been discussed under the heading of informing providers; and utilization review, practice profiling, detecting medical errors, and monitoring quality have been discussed under the heading of monitoring provider performance. The reason to reiterate them here is that these tools go beyond simply informing providers and monitoring their performance. They are actually rules which must be followed by providers, and activities that monitor whether providers are following these rules. It is these rules that collective purchasers use to participate in medical decisions. To avoid duplication, we provided a summary of those tools that have been discussed, and emphasize here two which have not been addressed: pre-authorization and second surgical opinion programmes.

Table 3 summarizes the evidence of effects of the policy tools which have been discussed earlier, and which can be used in participating in medical decision-making by collective purchasers. The use of medical practice guidelines has been proved to have a positive effect on allocative efficiency. Utilization review may have improved the mix of health interventions, but it may not be preferred because it cannot reduce overall cost and there is no evidence that it can improve health outcomes. There is still a lack of evidence on all other tools in terms of whether they improve allocative efficiency, either because the tool is not well designed (e.g., prescription formularies) or because implementation is still at an early stage and there are no rigorous evaluation studies.

Table 3. Policy tools used by collective purchasers for participating in medical decisions: evidence on improving allocative efficiency

POLICY TOOLS	EVIDENCE ON IMPROVING ALLOCATIVE EFFICIENCY	CAVEATS
Medical practice guidelines	There is substantial evidence that having providers guided by medical practice guidelines can not only control cost, but also improve health quality and outcomes, because it can reduce the provision of cost-ineffective interventions and increase the provision of cost-effective ones.	It is believed that the strength of effect depends on the degree to which the guidelines are based on cost-effectiveness information and on the extent to which guidelines are followed by providers.
Prescription formularies	There is no evidence on whether the use of formularies can increase the use of cost-effective drugs and decrease the use of cost-ineffective drugs. Evidence shows that it can reduce drug costs, but not the overall cost of health care.	The major reason for the failure is that the design of formularies is largely based on the cost of drugs rather than on cost-effectiveness information.
Utilization review	There is evidence that it can reduce the use of elective surgeries, hospital days and expensive drugs which are considered to be unnecessary or inappropriate. This suggests that utilization review may be able to improve the mix of interventions. However, the lack of evidence on its effect on improvement in quality and health outcomes and the increase in overall cost give rises to doubts about whether it can improve overall efficiency.	Case-by-case utilization review is a costly process. The fact that it can reduce utilization, but cannot decrease overall cost, and does not increase health outcomes suggests that it may not be a worthwhile practice.
Feedback	Evidence is limited on drug utilization. Face-to-face interactions with doctors on their deviated prescriptions can reduce drug cost, but it is not certain that it can reduce overall cost and improve the quality of care. There is a lack of evidence on whether it can reduce the use of cost-ineffective drugs and increase the use of cost-effective drugs.	Feedback based on information generated from performance monitoring on doctors' deviated provision behaviour seems to be necessary, but the impact of feedback needs to be further studied. Efforts should be made to find cost-effective ways of feedback.
Medical practice profiling	Although its practice is increasingly common, the effect of medical profiling is not well evaluated.Rigorous studies are yet to be done.	Profiling is preferred because it is less costly, less intrusive, and can identify systematic problems.
Detecting medical errors	The practice is at its early stage. The ways medical errors can be detected and reported need to be developed. Evaluation studies are yet to be done.	Because of the proven existence of large amounts of medical errors, detection of medical errors and finding ways to eliminate them will improve efficiency of health care.
Monitoring quality	Internal quality assurance programmes have a long history, but continuous, systematic and persistent quality monitoring is at its early stages. Evaluation studies are yet to be done.	Maximizing quality of care within a given budget constraint is at the heart of provider performance. It is suggested that quality evaluation should focus on process and outcomes, rather than structure, for the purpose of improving allocative efficiency.

Chapter 9. Towards a Policy Toolkit

The objectives of this section are to suggest a toolkit for improving allocative efficiency based on combination of the available tools, and to recommend future policy actions and research.

Before designing a policy toolkit, policy-makers must make a difficult political choice in addition to following the principles of tool development (see Annex). Given a budget constraint on resource allocation, should priority be given to covering the large economic risk of illnesses that usually need high cost interventions with only marginal health benefits, or should it be directed to financing an essential package of interventions which are cost-effective? The choice of the former can protect people from financial risk and prevent poverty resulting from the use of costly interventions, but it may lead to cost escalation and competition for resources to the detriment of essential and cost-effective interventions. The choice of the latter can assure access to the essential package of interventions by all, but does not protect people from the financial risk of costly interventions.

In case a country cannot do both, a compromise is that government takes responsibility for organizing and regulating the financing of an essential package of interventions (via either public or private financing or both), and leaves all other interventions to the private market. The results is that essential interventions are accessible for all, but the costly interventions are accessible only to those who are able and willing to pay for private health insurance or pay out-of-pocket at the point of services. This leads to a two-tier system — the rich enjoy more care and less financial risk; the poor enjoy less care and more risk. This compromise is far from perfect, but it is much better than the situation in which the rich are covered for everything and the poor are not covered for anything, because resources are allocated to cost-ineffective interventions, and the provision of cost-effective interventions is squeezed out.

Allocative efficiency is a function of many main factors: input mix for the production of interventions, supply of interventions, demand for interventions, doctors' behaviour, collective purchasers' behaviour, policy interventions in the health care market, etc. To improve allocative efficiency, policy tools that affect each of these main factors are needed.

Things become complicated, because each of the main factors is a function of many subfactors. For example, input mix depends on human resources (education policy, price for education, expected earnings after graduation); capital (capital investment policy, demand for care, price of capital input, expected rate of return); geographical balance (resource distribution policy, the demand for care in different areas, expected earnings in different areas), and similarly for the other main factors. In theory, policy development should be based on evidence of the effect of each of the subfactors on main factors, as well as the evidence of the effect of main factors on allocative efficiency. Such a policy development process will result in a suggested package of interventions, which should be tested simultaneously. In the real world, policy development has never gone this way because it is hard to determine valid cause–effect relations between dependent and independent variables because of measurement problems and the difficulty of conducting social experiments.

Fortunately, various policy tools have been developed and tested as solutions of specific problems related to allocative efficiency. A second-best choice is to review these tools, select from among them the potentially effective ones, and compose them into a toolkit. The use of

a toolkit to improve allocative efficiency is desirable because the combined effect of many policy tools should be greater than the sum of each tool's effect alone; the combination of tools reflects the real world situation where many policies exist for the achievement of the goals of health care systems; and sometimes it is difficult to separate the effect of one policy tool from that of another.

Almost all studies have tested the effect of one particular policy tool, and there is very little evidence on the effect of combined policy. While research is needed to test each of the possible tools, it would be a mistake to use only one tool to solve a problem, because a problem of the health system is usually an outcome of many factors.

Another argument for the development of a toolkit is that every tool has its strengths and weaknesses. Often, the success of one tool requires the success of another; one tool is used to overcome the shortcoming of another; or one tool helps to strengthen the effect of another. Sometimes the strength of the tools can be brought into full play, and the weaknesses limited, only if tools are combined.

Controlling intervention inputs

1) Budgeting input for public provision Governments are in a good position to improve allocative efficiency by wise use of their budgets for health, particularly when government funding for public provision accounts for a significant part of total health expenditure. A major problem is that government budgets are usually input-oriented without a clear vision of the types of health interventions provided. To improve allocative efficiency, government funding has to be output-oriented. The specification of health interventions which are relatively cost-effective is the key to improving allocative efficiency.

2) Public funding for cost-effective products Government should take an active role in financing medical products which are cost-effective and which producers are less willing to supply because of limited demand (e.g., vaccines). In countries where public production dominates, governments should provide sufficient funds for the production of these products; where private production dominates, government should provide subsidy.

3) Human resources planning The mix of health professionals, as well as their locations, are associated with the mix of health interventions. A cost-effective mix of health professionals is essential for the provision of a socially desirable mix of interventions. Human resources planning can start from the planning of medical education by controlling the types and the number of student enrolments, and can also provide incentives to motivate enrolments for health professionals in shortage, to encourage them to work in underserved areas, and to pay for their services relatively generously through fee structuring.

4) Health facility planning The types, quantity, size and location of health facilities are associated with the mix of health interventions provided. To improve allocative efficiency, health facilities should be planned so that a package of cost-effective interventions should be available for all, provided at the lowest possible level of administration, and cost-ineffective interventions are controlled to an acceptable minimum.

5) Capital planning The number of beds, the stock of major medical equipment, and the total capital investment can be brought under the control of governments. Overinvestment in capital can lead not only to cost-escalation, but to overprovision of capital-related services. A quota for the numbers of beds, and equipment per 1000 population should be used; major investment should be certified by government agencies; and capital cap should be used to control the overall spending on capital investment.

6) Technology regulation This can serve as a powerful tool to improve allocative efficiency by prohibiting cost-ineffective health interventions from entering the medical market. Effectiveness depends on the availability and validity of evidence on cost-effectiveness generated from technology assessment, and the strengthening of technology regulation. Existing technologies without proven cost-effectiveness should be rigorously assessed. Those which are not cost-effective and for which there are cost-effective substitutes should be eliminated from the market.

Managing intervention provision

7) Medical practice guidelines There is substantial evidence that the use of medical practice guidelines can reduce cost and improve health outcome, as well as change the mix of interventions. Although there is no direct evidence that it can improve allocative efficiency, the potential seems to be unshakeable. To strengthen the effectiveness of medical practice guidelines, two points are of key importance: the medical practice guidelines must be based on cost-effectiveness of alternative interventions; and they must actually be used by practitioners in their medical practice. The former relies on research for generating and disseminating information, and the latter depends on monitoring the use of the guidelines.

8) Prescription formularies These can induce practitioners to select cost-effective pharmaceutical interventions, and improve the interventions mix. To develop them into an effective tool to improve allocative efficiency, the formularies have to be based on cost-effectiveness information, rather than on the level of cost alone. Also the formularies should not be too rigid, and should allow for practitioners' decisions to reflect the special needs of patients.

9) Utilization review This can improve allocative efficiency if it is consistent with evidence-based medical practice guidelines. Case-by-case utilization review is costly and is not recommended. If it is based on the information generated from medical practice profiling and captures only the treatments and practitioners that fall outside the normal range of provision, and if it is based on sampling review as a part of provider performance monitoring, it may have a positive effect on allocative efficiency. It is more likely to be effective if the results are used to provide feedback to practitioners, and are linked to penalties and rewards.

10) Medical practice profiling This is an effective tool to monitor the provision behaviour of practitioners. The objective is to detect behaviour which has a likelihood of providing cost-ineffective interventions when cost-effective interventions are available. The effectiveness of this tool will be strengthened if it is combined with utilization review, feedback to providers, and penalties and rewards.

11) Feedback to providers This is based on the results of medical profiling, utilization review, and provider performance monitoring. Informing providers of their deviant practice and letting them double-check whether they are following medical practice guidelines can lead to improvement in their provision. The method of feedback is important for achieving the desired effect: it is most effective on an individual basis.

12) Pre-authorization This should target the interventions for which there is evidence of inappropriate provision, or which are costly and elective. The purpose is to restrict the use of costly interventions when less costly interventions with similar health effects are available. Appropriate use of this tool can reduce cost, maintain the quality of care, and improve allocative efficiency.

13) Second surgical opinion programme This should target the surgical procedures which are either costly or risky, and where the health effect is often ambiguous. The procedures included must be clearly defined, and the programme should be compulsory. This is expected to reduce the number of costly surgeries with uncertain health benefit, and thus to improve allocative efficiency.

14) Regulating intervention availability The availability of essential health interventions, which includes essential drugs and services which are considered cost-effective, is a requirement for promoting their use. When these interventions are not available either because of lack of resources or because of lack of motivation to provide them, governments can regulate the types of drugs that must be available at different levels of health institutions, and the structure and capacity of providers at different levels to deliver the essential interventions. Availability is essential, but not sufficient. The effect of this tool also depends on whether these essential interventions are covered by health care financing schemes.

15) Educating providers The knowledge base of practitioners, who are the real decision-makers for allocating health interventions, is important for them to play a better role in the allocation of social resources. Traditionally, practitioners are responsible for the health of individuals they treat, and recommend interventions that are effective without considering their cost. Practitioners should be turned into socially responsible providers, educated so that they are aware that resources are limited and the provision of cost-ineffective interventions will be at the sacrifice of cost-effective ones. They need a good grasp of medical knowledge as well as of the cost-effectiveness of alternative interventions, so they can take on the role of improving allocative efficiency by considering cost-effectiveness of alternative interventions when making medical decisions. This tool can be strengthened by rewarding those who do so.

16) Monitoring provider performance This tool can be used in conjunction with many other tools. It is a prerequisite of performance-related pay. The results of performance monitoring can be disseminated to consumers for empowering their choices; they can also be used to provide feedback to providers for improving their performance. The use of provider incentives that reduce provision has to be combined with monitoring performance to prevent underprovision of necessary interventions. The focus of performance should include quality (process and outcome) and the mix of interventions within a budget constraint. The effect of this tool on allocative efficiency depends on whether the overall efficiency of the mix of interventions is taken as a performance measure.

Purchasing interventions

17) Fee-for-service If appropriately designed and implemented, this can serve to improve allocative efficiency. FFS can be used to pay for essential interventions (e.g., immunizations), to motivate providers to provide more of these interventions. By structuring the fee schedule in such a way that cost-effective interventions are reimbursed more generously, it can improve allocative efficiency. Incentives for overprovision of cost-ineffective interventions can be controlled through joint use of other payment methods (e.g., capitation) and a global budget. To strengthen the effectiveness of this tool, cost-effective interventions should be included in insurance benefits, so that the increased fee for cost-effective interventions does not become a barrierfor users.

18) Capitation This is a tool to control costs as well as to improve allocative efficiency, because it motivates providers to provide less costly, preventive, and cost-effective interventions. The effect can be strengthened by introducing FFS payment for specified essential interventions. Still better effects can be achieved if capitation payment includes a wide range of interventions; the performance of providers is monitored; and the results of monitoring are used to adjust the level of capitation payment. There should also be competition among providers for consumer registration, and capitation payment should be risk-adjusted.

19) Salary This can be used to improve allocative efficiency, because it provides no incentive for either overprovision or underprovision. To strengthen the effect of this tool, two things are important. One is doctors' knowledge of cost-effectiveness of alternative interventions, which must be strengthened through both education and the use of medical practice guidelines. Another is to prevent low morale through, for example, salary or promotion based on performance, or bonus payment based on performance on top of the salary.

20) Case payment This can be used to improve allocative efficiency if provider performance monitoring prevents providers from underprovision of necessary interventions. Case fees should be structured to allow more generous payment for the diagnoses for which cost-effective treatments are available; case payment should be as inclusive as possible, and implemented for all payers with a global budget to cap the overall spending for health care.

21) Global budget This is a powerful tool to control costs. To improve allocative efficiency, policy-makers and providers must be educated to use the budget wisely (providing more cost-effective interventions and fewer cost-ineffective ones), with incentives for providers to do so. The implementation of a global budget requires an integrated payer system, and a full global budget is more effective than a partial one. Almost all other tools can be combined with global budgets for improving allocative efficiency.

22) Performance-related pay The effectiveness of this tool depends on how performance is measured. It can be effective for improving allocative efficiency, if both quality and cost are considered in designing the PRP for individual doctors, and if the overall efficiency of the mix of interventions is considered in designing the PRP for institutionalized providers.

23) Fee structuring This is expected to be a powerful tool to improve allocative efficiency. With FFS, among numerous health interventions, the prices for cost-effective interventions can be set at relatively high levels to encourage the provision of these interventions; and the prices for cost-ineffective interventions can be set at relatively low levels to discourage their provision.

With case payment, among various diagnoses, the payment for those cases with cost-effective treatments or with greater need for hospitalization can be paid more generously to encourage admissions; and payment for those cases with no cost-effective treatment, or with less need for hospitalization can be paid less generously to discourage their admissions. Correction of a distorted fee schedule is expected to have a positive effect on allocative efficiency. Fee structuring may not work for health systems where out-of-pocket payment dominates.

24) Removing incentive for prescribing When the monetary value of a doctor's prescription is linked to his or her income, doctors are motivated to prescribe more and costlier drugs, which is a departure from allocative efficiency. There are three options for removing this incentive: organizational separation of dispensing from prescribing; reducing the mark-up rate for drugs; and reimbursing dispensing services based on the number of patients served rather than on monetary value or the number of drugs per prescription.

25) Payer integration This is an arrangement that integrates numerous flows of funds into one or several major flows. It is a prerequisite for the implementation of global budgets through which the allocation of resources among health interventions can be controlled. Payer integration by itself does not directly improve allocative efficiency, but it is an important condition for many other tools related to providers' financial motivation.

26) Motivating users This tool includes user fees and cost sharing, the essence of which are to increase the patient's financial responsibility for the use of health interventions. To improve allocative efficiency, users' financial responsibility should increase as the cost-effectiveness of health interventions declines. The incentive for users should be designed to allow access to basic and highly cost-effective interventions for all.

27) Informing users Government takes an active role to inform consumers and to empower their choices of providers as well as of interventions. For the former, governments need to collect information (price, cost, quality of care, etc.) from providers and disseminate this to the general public in a cost-effective manner. For the latter, governments need to regulate providers to provide necessary information (costs and effectiveness of alternative interventions) at the point of medical consultation so the patients or their surrogates can participate in medical decisions and make better choices of interventions.

28) Regulating collective purchasers With this tool the coverage of essential health interventions by collective purchasers is regulated so that utilization of cost-effective interventions can be promoted. Collective purchasers need to be regulated also on their use of financial incentives for providers, to prevent misuse of incentives for the benefit of purchasers to the detriment of consumers. These regulatory activities help to ensure that inexpensive cost-effective interventions are used fully, and that costly cost-effective interventions are not underutilized.

Several tools which were reviewed earlier are not included here; and several new tools have been added to the list. Strengthening medical education, strengthening health facilities and training grant incentives are included in human resources planning and health facility planning. Certificate-of-Need, capital cap and capital moratorium are included in capital planning. Daily payment is excluded because it is not expected to be used to improve allocative efficiency. Line budgets (separate budgets for different costs) for institutions are not thought of as policy tools, but can be used to allocate government budget if the budget line specifies the type of interventions (separate budgets for different interventions). It is not regarded as an independ-

ent tool because it is included in budgeting input for public provision. Rate setting is removed from the list because it is the fee structure that is relevant to allocative efficiency, and that is included in fee structuring. User fees and cost sharing are not regarded as separate tools; they are included under the heading of motivating users. Rationing is removed because it is included in user fees and cost sharing, and also because it has a negative connotation which may not be acceptable politically. Essential drug lists and essential packages of interventions are included in budgeting input for public provision, fee structuring, motivating users, and regulating collective purchasers. Managed care is removed because it is not an independent tool. Medical practice profiling, pre-authorization, second surgical opinion programme, educating practitioners, monitoring provider performance, informing users and regulating collective purchasers are added to the list as independent tools. Removing incentive for prescribing is considered a substitute for separating dispensing from prescribing, because it is the incentive that matters.

Toolkit construction

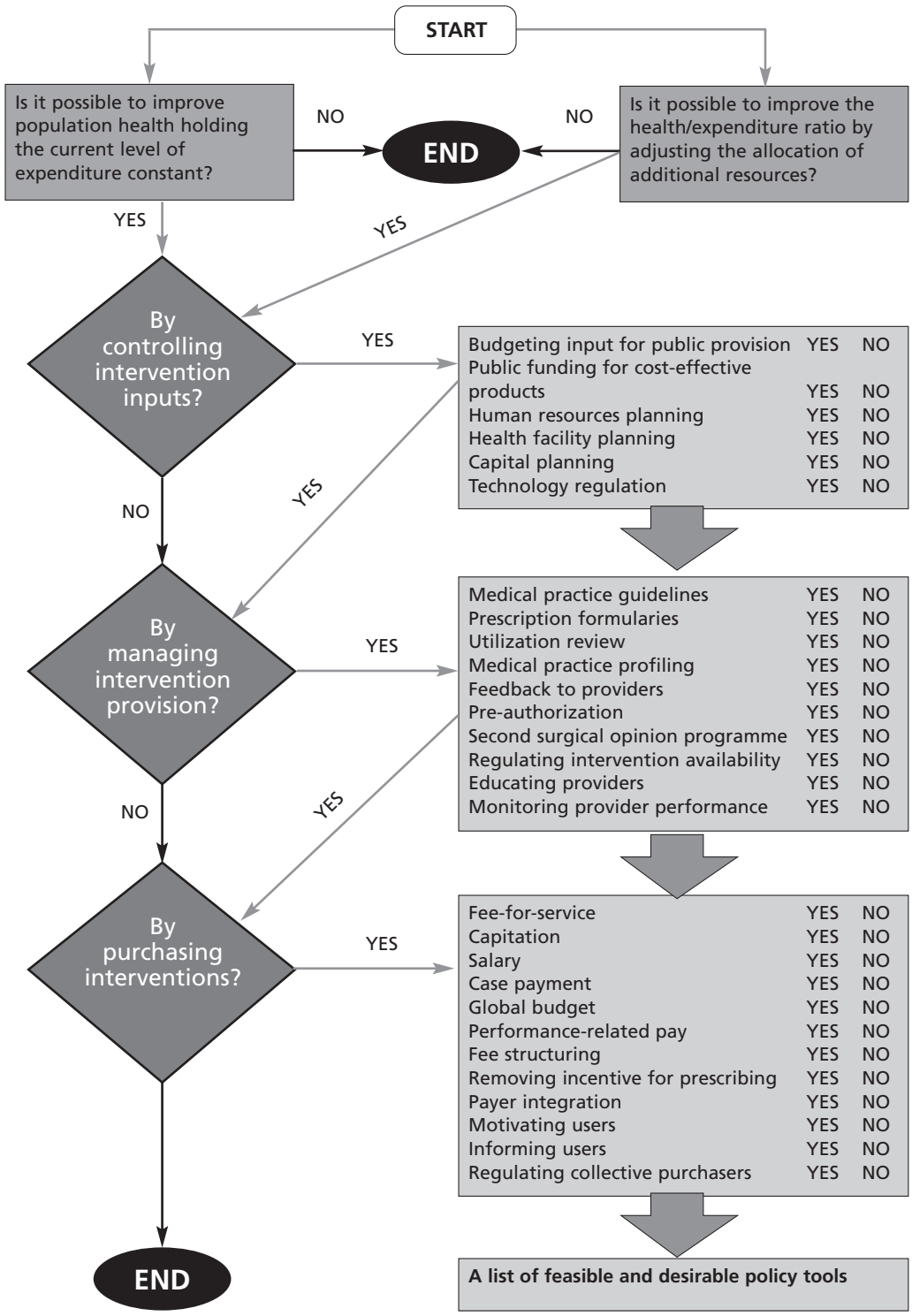
It is difficult to recommend universally applicable specific tools, because the economic, political, and health background varies among countries. However, this list of possible tools, for policy-makers in different countries to choose from, can be used to construct a toolkit. In constructing a toolkit, policy-makers should first consider what are the major problems of allocative inefficiency. This involves a thorough situation analysis which must be conducted before constructing a toolkit.

After the problems are identified and the decisions on priorities made, the next step is to follow a flowchart (Figure 7) for choosing tools to be included in the kit. The process starts by asking two parallel questions. First, is it possible to improve population health holding the current level of expenditure constant? Second, is it possible to improve the health/expenditure ratio by adjusting the allocation of additional resources? If the answer to both these questions is "No", which is usually not the case, there is no need to develop tools to improve allocative efficiency. If the answer to one of the two questions is "Yes", the selection process begins by asking three sequential questions which are indicated by the three diamonds in Figure 7. If the answer for each of them is "Yes", the process goes on to the selection of specific tools. The sequence goes from top to the bottom, ending with a list of selected policy tools.

Before the selected tools are put to use, two tests are needed. First, the tools selected must be evaluated against the five principles proposed in Annex 1 (efficiency, evidence-based tool development, covering all policy anchoring points, covering all stakeholders, and feasibility). There is no excuse for keeping any tool which is not likely to have a positive effect on efficiency. Because evidence is still hardly available for almost all, the kit should be tested before launching large-scale implementation.

Second, some policy tools are mutually dependent, and some of them are mutually complementary: simultaneous implementation is more effective than the summed effect of their implementation separately. We suggest checking the conditions for the success of each of the tools as stated in the list of possible tools in this section, so that the package of tools can supplement or complement each other. For example, case payment and capitation need intensive monitoring of provider performance; FFS payment needs payer integration, a global budget, and appropriate structure of the fee schedule; utilization review should be based on the implementation on of medical practice guidelines, and so on.

Figure 7. A flowchart for the selection of policy tools in a toolkit for improving allocative efficiency



Chapter 10. Recommendations for Policy and Research

Capacity strengthening

Surely every government is willing to achieve better health for its population with the current level of health expenditure. The problem is that many governments face constraints for achieving this. These include the limitation of economic capacity in the poorest countries, which prohibits the provision of access to basic health interventions; the political constraints which hinder the appropriate flow of resources; and the technical constraints which become major barriers for the government to make and implement good policies. While the economic and political constraints should not be ignored in the design and implementation of policies, the focus here is on technical capacities which are more likely to be strengthened through both national and international efforts.

Stimulating demand for evidence The lack of demand for evidence for health policy is a major reason for the underdevelopment of health policy and system research. In many countries, policies are based on dictatorship from the top, or on compromise among different opinions within the leadership. Because of lack of demand, academic experts are not very willing to provide evidence based on applied research. In addition, researchers may worry that they will run into trouble if they provide research results that question or contradict the decisions from the top. Although the culture of evidence-based policy-making has been nurtured over the past decade, it is still far less strong than is needed. Besides the need for political reforms that promote democracy, policy-makers should be convinced that policies based on evidence are better than those based only on opinions or theories. It is expected that advocacy will increase the demand for evidence, and will foster health policy and systems research.

Building a core team of experts The absolute majority of the publications found and used in this effort for searching for policy tools are from high-income countries. Of those related to health policy in low- and middle-income countries, most are authored by researchers from high-income countries. Although there might be other reasons for the lack of publications from low- and middle-income countries, the lack of qualified researchers is a major one. Indeed, in many low- and middle-income countries there are no more than a handful of experts; and many of them work on proving the positive effect of government policies, which are likely to be biased, rather than providing scientific evaluation and evidence. There is an urgent need to build a core team of experts who can provide technical and scientific support to both local and central government for making better health policies.

Developing information systems Providing evidence for policy is not a one-off effort. Because policy-making, implementation, evaluation and modification are a continuous process, they need continuous information support. Information systems which provide accurate, continuous, and timely information and evidence are a prerequisite for evidence-based policy-making. In broad terms, the system should provide raw data that can be analysed to produce information as well as published information, and national as well as international information, so that one country can borrow from the experience of another.

Strengthening the use of information It is often the case that government-sponsored data are available, but they are not accessible to researchers and policy-makers. This leads to

underutilization of data which have a large potential to generate benefits for public policy. Publicly sponsored databases are public goods which should be used fully by those who work for policy research and policy-making. To maximize the benefit from the existing databases, governments should ensure that there is a free flow and use of information, free publication of results, and free expression of opinions. Unless existing databases are fully used and translated into information and evidence, they are of little value.

Building capacity for informing policy-makers How to connect research results, conclusions and recommendations with policy-making is a challenge for many countries. Policy-makers are often confused by complex statistics and results, which provide little information for policy. To make better use of research results and to inform policy-makers, results should be presented in a user-friendly way, and interpreted with a focus on policy implications. Collaboration between policy-makers and researchers can promote mutual understanding, and promote policy-oriented researches and the adoption of recommendation by policy-makers; and a formal and regular channel for dialogue (e.g., a policy advisory committee for the government) can foster the translation of research into policy actions.

Supporting health policy and systems research It is often said that health policy and systems development is important because it is a cross-cutting area. However, just because of this, it becomes neglected because there is rarely a department in either central or local governments which is specifically responsible for it, and collaboration between different departments within the central or local health authorities is often difficult. In addition, health policy and systems research often does not yield immediate and visible results. Because of this, government is often reluctant to provide funding. It is increasingly recognized that health policy and systems are at the heart of the health sector; and that a little improvement in policy may lead to a big change in system performance.

Policy-oriented actions

Several specific technical barriers exist to evaluating whether a policy tool has an effect on allocative efficiency. Besides the general suggestions stated above, several specific policy-oriented recommendations that target these barriers are suggested for researchers and policy-makers at national level as well as the international community.

At national level, the following actions need to be taken: prioritizing diseases and risk factors; cost-effectiveness analysis of health interventions; situation analysis of allocative efficiency; and developing policy tools.

Prioritizing diseases and risk factors This is an old suggestion, but it should be continuously and strongly recommended, because very few countries have done so. At national level, priority of resource allocation should depend on the size of the burden of disease and the availability of cost-effective interventions. As the first step, in order to improve allocative efficiency, the diseases and risk factors should be ranked according to the size of the burden on the population, and resources should be directed to those diseases and risk factors which bring about the greatest burden and for which cost-effective interventions are available. The method for estimating the burden of diseases is available (4), and whether this can be done depends on the availability of surveillance data on the causes of death, diseases, risk factors and injuries.

Cost-effectiveness analysis of health interventions The information on cost-effectiveness of interventions has two uses. One is for prioritizing the interventions to deal with priority diseases and risk factors; another is for evaluation of the impact of policy interventions on allocative efficiency. For the former, in cost-effectiveness analyses, health interventions must be matched with diseases and risk factors. For the latter, it is known that the implementation of some of the tools has led to changes in intervention mix (e.g., increase in use of essential drugs and decrease in use of non-essential drugs; decrease in inpatient services and increase in outpatient services; increase in preventive interventions and decrease in elective surgeries). No conclusion can be reached, however, on whether these changes mean an improvement in allocative efficiency because of the lack of information on the cost-effectiveness of these interventions.

Situation analysis of allocative efficiency Situation analysis is based on and goes beyond prioritizing health conditions and health interventions. The major tasks of situation analysis are to examine whether the current allocation of resources matches the priority health conditions and cost-effective interventions; to analyse the existence and the size of the problems of allocative inefficiency; and to explain the reasons for the problems. These activities are essential both for advocacy and for the development of policies for improving the efficient use of health resources.

Developing policy tools Policy tool development should be based on situation analysis, so that tools are developed according to the reasons for major problems of allocative inefficiency. Governments are urged to consider the recommended tools listed in Figure 7, and select and adapt these tools according to their individual situations. Tools can be developed and tested individually, but it is expected that the implementation of a toolkit is more effective. The evaluation of the tools should capture their effect on both cost and health outcomes, and should quantify their effect on the overall efficiency of the mix of health interventions.

At international level, there is a lot to do to provide technical support for countries which are facing the same technical barriers. To improve the efficiency of efforts, and also to offset the lack of expertise and resources in most low- and middle-income countries, some of the work should be done at international level, so that the methods and information generated can be generalized and shared by all countries, at least in the following areas.

An international repository of information on the cost-effectiveness of health interventions Economic evaluations of health interventions have been done in many countries, but these are not conducted in a systematic manner. There are three problems with the past work. First, economic evaluations in a single country cannot cover all possible interventions and match interventions to diseases, injuries and risk factors. Second, there is no standardized and uniform measure for effectiveness — or the current measure (DALYs) is not used — to allow for comparisons across interventions. Third, the cost measure is mostly based on local currencies and prices. It is suggested that an international repository should be set up to work out a systematic plan for cost-effectiveness analysis of health interventions; standardize both effectiveness and cost measures; gather cost-effectiveness information from all countries; and synthesize, standardize, and disseminate information to policy-makers and researchers. DALYs can be used as a standard measure for effectiveness, and purchasing power parity can be used as a standard cost measure. WHO-CHOICE provides cost-effective information to assist policy-makers in choosing health interventions, and represents an effort towards establishing such a repository.

An international centre for benchmarking practice Medical practice guidelines have been proved to have an effect on improving quality, reducing costs and increasing allocative efficiency. However, numerous medical practice guidelines are only available in some OECD countries, and are not widely shared. There is a need to set up an international centre for benchmarking medical practice, whose functions are to work out a systematic plan for development of medical practice guidelines; collect existing medical practice guidelines and share them among countries; and modify these guidelines so that they can be used by low- and middle-income countries.

Measuring provider performance Provider performance is the key to health system performance. Efforts should be made to improve the performance of health care providers. While performance-related pay and performance monitoring are becoming popular, there is no consensus on how provider performance should be measured. International efforts should be made to develop provider performance measures in line with the measures of health system performance. Measures developed by purchasers and institutionalized providers are for their own goals, independently of the goals of health systems.

The design and implementation of a toolkit to improve allocative efficiency as well as health system performance is not an easy task. It requires both technical expertise and political support. It requires international efforts as well as efforts at country level where the health system is organized. It demands continuous efforts by both policy researchers and policy-makers. It requires joint efforts of all parties within and outside of the health sector. A better health system is not something to talk about, but something to create. If policy-makers decide to do something, there is a hope to achieve better health system performance and ultimately health outcomes with the money spent. The joint and persistent efforts of policy-makers and researchers, both national and international, will be rewarded by better health systems, to be enjoyed by all the people of the world.

Annex: Principles for Tool Development

For the development of a useful toolkit, several principles have to be followed. First, the toolkit must be able to improve allocative efficiency. Second, development must be based on evidence of effects on health, cost, and the mix of health interventions. Third, the kit must include policy tools that affect all three issues of controlling inputs, managing provision, and purchasing interventions. Fourth, the kit must apply to all relevant parties responsible for resource allocation. Fifth, the toolkit must be technically and politically feasible.

Efficiency is the first principle for the development of a toolkit to improve allocative efficiency. Suppose the cost-effectiveness, potential use, and total cost of all health interventions in a country are known. The interventions can be ranked according to their cost-effectiveness; and the accumulated cost can be estimated based on the ranking of interventions. The efficiency principle tells us that the provision of health interventions should start from the most cost-effective health interventions and move towards the least cost-effective interventions until the budget runs out, or the marginal health benefit becomes worth no more than the marginal cost.

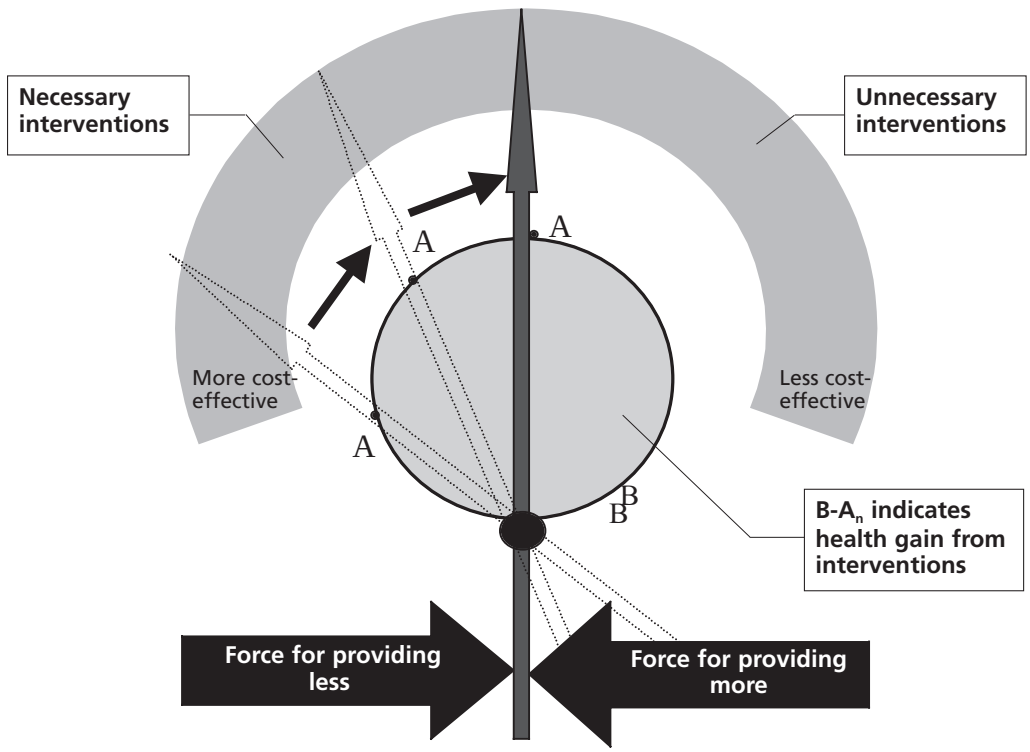
This principle is illustrated in Figure A1, which indicates the hypothetical effect of policy tools on the provision of health interventions, cost, and health. The shaded arc indicates the accumulated cost of all available health interventions ranked from the most cost-effective on the left to the least cost-effective on the right. The pointer indicates the mix of health interventions and health gain pivoting on point B. Moving the pointer to the right is equivalent to going down a list of interventions from more to less cost-effective. The pointer is positioned by policy tools that affect provision and utilization.

If provision starts from the left (the most cost-effective interventions), the total health gain (B-A) increases with the increase in provision of interventions. As the pointer moves from A1 to A2 and then to A3, the incremental health gain increases at a decreasing rate, and the marginal cost per unit of health gain increases at an increasing rate. As a result, the more health interventions are provided along the arc, the lower the overall efficiency of the health care system. It can also be seen that when the provision passes the optimal point A3, the level of cost continues to increase, but the marginal health gain becomes negative, because the care provided is inappropriate, unnecessary, or even harmful. The implication is that, if health resources are sufficient, the level of provision should not pass A3 where total health gain is maximized — and the gain is assumed to be worth the total cost to society; and that if there is a budget constraint, provision of course stops where the budget runs out.

In the real world, provision does not start from the left and move to the right. In addition, this theoretical argument is based on the assumption that the right interventions are delivered to the right people. In fact, cost-effective interventions are sometimes wrongly used by consumers; in which case, they cannot produce a desirable effect on health. Nonetheless, these arguments are useful since they imply that it would be a mistake (even it is unavoidable) to provide health interventions further to the right unless interventions to the left are already provided. They also imply that health interventions beyond the optimum A3 should not be provided in any event.

Where the provision starts, whether there are more cost-effective interventions left behind, and whether there is provision of interventions beyond the optimum are all affected by policies which may affect the provision process in both directions — towards providing more and towards providing less. A well-designed toolkit is expected to combine the two forces towards a better mix of health interventions.

Figure A1. The hypothetical effect of policy tools on the provision of health interventions and health outcomes



Evidence-based development is the second prerequisite for success of the toolkit. Most of the policy tools reviewed here were designed and implemented, based not on evidence, but on theoretical predictions and beliefs. Fortunately, the implementation of these tools gave rise to social experiments in the real world; and experiences of success and failure have been generated. Now may be the right time to suggest a toolkit, even though evidence on current policy tools is far from sufficient because very few tools were designed to improve allocative efficiency, and their indirect effects have not been rigorously evaluated.

The third important principle for the development of a toolkit is that they cover all *policy anchoring points*, which affect the control of intervention inputs; management of provision; and purchasing of interventions.

There are two caveats related to this. First, managing intervention provision is crucial because it is the providers who decide what interventions are provided to what people. Tools designed for controlling intervention inputs and purchasing interventions try to influence the medical

decision process so that the care provided is cost-effective. Second, it may be argued that perfectly informed purchasing can determine the provision of interventions; and the provision of interventions will then determine inputs. In that case, policy tools should be focused on purchasing interventions. However, perfectly informed purchasing is unlikely to exist. Although better-informed consumers are more likely to make rational choices of both providers and interventions, it is almost impossible for them to know as much as providers. Although collective purchasers have more power to influence providers' medical decisions, they are not likely to be perfect agents for consumers because of the information gap between purchasers and providers, and also because of purchasers' self-interests that may conflict with the interest of the consumers. Thus, focusing on one policy and neglecting others may not be the right choice.

Participation by all stakeholders is necessary to have the toolkit work well. The health production process involves government (policy-makers), providers, collective purchasers, and consumers. A well-designed toolkit should have influence on all of them so that there are joint efforts and concerted actions. As indicated in Table A1, each actor in health care systems performs its own actions, through which the different parties interact. A toolkit should be aimed at all parties and their actions, so as to promote the better use of limited health resources.

Table A1. Actors and their actions in the process of health production

GOVERNMENTS	PURCHASERS	PROVIDERS	CONSUMERS
Regulating	Financing	Producing care	Paying premiums
Rule setting	Budgeting	Financing providers	Paying fees
Policy-making	Purchasing	Allocating intervention	Using interventions
Informing	Paying	Informing	Choosing interventions
Monitoring	Monitoring		Choosing providers
Budgeting	Informing		Choosing insurers

Feasibility is the last principle for the development of a toolkit. A toolkit is not just something to show and read, but something to be implemented. The implementation of a toolkit will be harder than implementation of just one tool.

Technically, the following aspects of feasibility should be considered:

- Tools included in the kit should complement or supplement each other, and should not be in conflict.
- Designers of the toolkit should predict the possible unintended negative effect and tools that could be introduced to prevent those effects should be included. Often, policy-makers are very eager for a policy tool because they expect the intended positive effect, and overlook the possible unintended negative effect. If precautionary measures are not put in place against the unintended effect, it may outweigh the intended effect.

- Implementation involves considerable cost. The designers should take full consideration of the financial capacity to implement the toolkit.
- Technical capacity for developing and implementing a toolkit is of key importance. It includes expertise, as well as technical infrastructure (e.g., an information system). If the technical capacity is lacking, the right policy is to build it prior to implementation of a toolkit.
- Implementation is a continuous process. Once it is implemented, the programme must have continuous resource input. Relevant measurements must be available to evaluate the effectiveness of the programme, and to adjust it based on the results of implementation.
- The toolkit should be designed to adapt to existing and predicted problems, and economic, cultural and political situations. There is no universal toolkit which is suitable to every country. The toolkit suggested here is expected to provide a base for countries to design their own toolkits, suitable to their own situations.

Politically, the following aspects should be considered:

- Implementation requires real and full collaboration and coordination among different parties, and among different stakeholders within each party (e.g., different departments of the central government). The implementation of a toolkit is not a political football to be kicked from one player to another, but a policy which should be jointly accomplished.
- Although the implementation of a toolkit involves different parties, it always involves an active and strong government, which can provide good governance or stewardship, and is able and willing to accomplish the defined implementation agenda. The change of governing parties should not affect the implementation of the toolkit (since abandonment of implementation involves a high level of costs). Corrupt governments are not expected to implement a toolkit well.
- Any health care reform involves redistribution of interests; so does the implementation of these policy tools. The intended effect of policy implementation is to transfer more resources into population health, rather than transfer money from the pocket of one party to that of another.[1] For the former, commitment by all parties is easily agreed; but for the latter, conflicts of interests will arise. The toolkit must be politically acceptable to all parties. Strong rejection by any party will negatively affect implementation or the effectiveness of the policy toolkit.

In decreasing order of importance, there are several criteria for the selection of the tools. First, there is evidence that the tool has an effect on the mix of interventions. Second, there is evidence that the tool has an effect on cost containment and also the potential to improve health outcome. Third, if there is no evidence from evaluation, it is theoretically sound to expect that the tool can have an effect on the mix of interventions. Fourth, there is neither evidence nor theory in favour of the tool, but it forms a condition for the success of other tools. Some tools that do not meet even the least restrictive of these criteria are dropped, some new tools are added, and some selections are composed of several tools.

[1] This is somewhat the case in the United States where the implementation of managed care transfers money from providers to insurers without evidence of improvement in population health. For example, under unmanaged care, the doctors and hospitals made more money (e.g., the head of the Hospital Corporation of American made US$ 128 million per year before managed care (Navarro 1994) (402); with managed care the earning is transferred to managed care organizations (e.g., the biggest pay to the chiefs of managed care organizations ranged from US$ 4 million to US$ 1 million per year) (403).

Reference List

(1) World Health Organization. *The World Health Report 2000 — Health systems: improving performance.* Geneva, World Health Organization, 2000.

(2) Hawkins J, Allen R, eds. Oxford Encyclopedic English Dictionary. Oxford, Clarendon Press, 1991.

(3) Muir Gray JA. Evidence-based health care: how to make policy and management decisions. London, Churchill Livingstone, 1997.

(4) Murray CJL, Lopez AD. Evidence-based health policy: lessons from the global burden of disease study. Science, 1996, 274:740-743.

(5) Ham C. Evidence-based policy making (editorial). British Medical Journal, 1995, 310:71-72.

(6) Walt G. Health policy: An introduction to process and power. Johannesburg, Witwatesrand University Press, 1994.

(7) Meekison WG, Mackenzie CJG, Fraser B. Critical requirements in regional health planning. Canadian Journal of Public Health, 1975, 66:299-306.

(8) Angus DE, Auer L, Cloutier JE et al. Cost-effectiveness in ambulatory care: alternative approaches. Journal of Ambulatory Care Management, 1995, 18(1):29-38.

(9) Battista RN, Banta D, Jonsson E et al. Lessons from the eight countries. Health Policy, 1994, 30:397-421.

(10) Beadle CE. Addressing health care and health care costs at their roots. Preventive Medicine, 1996, 25:66-67.

(11) Segal L. Health funding: the nature of distortions and implications for the health service mix. Australian and New Zealand Journal of Public Health, 1998, 22(2):271-273.

(12) World Bank. World Development Report 1993 — Investing in health. Oxford, Oxford University Press, 1993.

(13) Wennberg JE, Gittelsohn A. Small area variations in health care delivery. Science, 1973, 182:1102-1108.

(14) Lipsitz D, Nagler H. A physician incentive compensation program in a staff model HMO. HMO Practice, 1993, 7(2):82-87.

(15) Winslow CM, Kosecoff J, Chassin M et al. The appropriateness of performing coronary artery bypass surgery. Journal of the American Medical Association, 1988, 260:505-509.

(16) Greenspan AM, Kay HR, Berger BC et al. Incidence of unwarranted implantation of permanent cardiac pacemakers in a large medical population. New England Journal of Medicine, 1988, 318:158-163.

(17) Restuccia JD, Kreger BE, Payne SM et al. Factors affecting appropriateness of hospital use in Massachuesetts. Health Care Financing Review, 1986, 8:47-54.

(18) Winslow CM, Solomon DH, Chassin M et al. The appropriateness of carotid endarterectomy. New England Journal of Medicine, 1988, 318:721-727.

(19) Schulenbury JM. Germany: solidarity at a price. Journal of Health Politics, Policy and Law, 1992, 17(4):715-729.

(20) Stoddart GL, Barer ML. Toward integrated medical resource policies for Canada:11. Improving effectiveness and efficiency. Canadian Medical Association Journal, 1992, 147(11):1653-1660.

(21) DeCoster C, Roos NP, Carriere KC. Peterson S. Inappropriate hospital use by patients receiving care for medical conditions: targeting utilisation review. Canadian Medical Association Journal, 1997, 157(7):889-896.

(22) Grand AL, Hogerzeil HV, Haaijer-Ruskamp FM. Intervention research in rational use of drugs: a review. Health Policy and Planning, 1999, 14(2):89-102.

(23) Hsiao WC. Abnormal economics in the health sector. Health Policy, 1995, 32(1-3):125-139.

(24) Enthoven A. The history and principles of managed competition. Health (Supplement), 1993, 1993a:24-48.

(25) Enthoven A. Market reform and universal coverage: avoid market failure. Health Care Strategy Management, 1993, 11(2):11-14.

(26) Fielding JE, Rice T. Can managed competition solve the problems of market failure? Health Affairs, 1993, 12:216-228.

(27) Bridgman RG. Regional planning of health care facilities and regional collaboration between health care institutions. World Hospital, 1975, Spring/Summer:65-67.

(28) World Health Organization. Role of hospitals in programmes of community health protection. Technical Report Series No 122 (WHO). Geneva, World Health Organization, 1957. http://whqlibdoc.who.int/trs/WHO_TRS_122.pdf

(29) Altman DE, Morgan DH. The role of state and local government in health. Health Affairs, 1983, 2(4):7-31.

(30) Walker C. An overview of the role of government in the organisation and provision of health services in Japan. Australian Health Review, 1996, 19(2):75-93.

(31) Mills A. Improving the efficiency of public health services in low and middle income countries: Bureaucratic versus market approaches. In: Colclough C, ed. Marketing education and health in low and middle income countries: miracle or mirage? Oxford, Clarendon Press, 1997.

(32) OECD. Health Data. Paris, Organisation for Economic Co-operation and Development, 1998.

(33) Birdsall N, James E. Health government and the poor: the case for the private sector. In: Gribble JN, Preston SH, eds. The epidemiological transition: policy and planning implications for low and middle income countries. Washington DC, National Academy Press, 1993.

(34) American Medical Association. Financial support of medical schools by the federal government. Journal of Medical Education, 1969, 44:996-997.

(35) Feldstein M. A preliminary evaluation of federal dental manpower subsidy programs. Inquiry, 1974, XI:196-206.

(36) McEldowney RP, Berry A. Physician supply and distribution in the USA. Journal of Management in Medicine, 1995, 9(5):68-74.

(37) Dunne MR, Miller RG. US graduate medical education, 1996-1997. JAMA, 1997, 278(9):750-754.

(38) Clark LJ, Field MJ, Koontz TL et al. The impact of Hill-Burton: an analysis of hospital bed and physician distribution in the United States 1950-1970. Medical Care, 1980, XVII (5):532-547.

(39) Dowell MA. Hill-Burton: the unfulfilled promise. Journal of Health Politics, Policy and Law, 1987, 12(1):153-175.

(40) Brown LD. The proper boundary of the role of government. Bulletin of the New York Academy of Medicine, 1986, 62(1):15-26.

(41) Leape LL. Unnecessary surgery. Health Service Research, 1989, 24(3):351-407.

(42) Tabbush V, Swanson G. Changing paradigms in medical payment. Archive of Internal Medicine, 1996, 156:357-360.

(43) Mullan F. The muscular Samaritan: the National Health Corps in the new century. Health Affairs, 1999, 18(2):168-175.

(44) Pathman DE, Taylor DH, Konrad TR et al. The National Health Service Corps experience for rural physicians in the late 1980s. JAMA, 1994, 272(12):1341-1348.

(45) Neushul P. Science, government, and mass production of Penicillin. Journal of the History of Medicine, 1993, 48:371-394.

(46) Beaulnes A. Vaccine development and production in Canada: a challenge in search of a policy. Canadian Medical Journal, 1981, 124:118-120.

(47) Simpson JB. State certificate-of-need program: the current status. American Journal of Public Health, 1985, 75(10):1225-1229.

(48) Hellinger FJ. The effect of Certificate-of-Need legislation on hospital investment. Inquiry, 1976, June 13:187-193.

(49) Salkever DS, Bice TW. The impact of certificate-of-need controls on hospital investment. Milbank Memorial Fund Quarterly Health and Society, 1976, Spring 54(2):185-214.

(50) Howell JR. Evaluating the impact of certificate-of-need regulation using measures of ultimate outcome: some cautions from experience in Massachusetts. Health Service Research, 1984, 19(5):587-613.

(51) Sloan FA. Controlling health expenditures: lessons learned from Certificate of Need Programs. In: Sloan FA, ed. Cost, quality, and access in health care. San Francisco, Jossey-Bass, 1988.

(52) Leeds HS. Certificate of need: up for revision. Health System Review, 1996, January/February:26-29.

(53) Harrington C, Swan J. State Medicaid nursing home policies, utilisation and expenditures. Inquiry, 1987, Summer 24(2):157-172.

(54) Swan J, Harrington C. Certificate of need and nursing home bed capacity in states. Journal of Health and Social Policy, 1990, 2(2):87-105.

(55) Moore J. Certificate of need: Gone in many states but not dead yet. Modern Health Care, 1997, August 11:32-33,36.

(56) Harrington C, Swan J, Nymon J et al. The effect of certificate of need and moratoria policy on change in nursing home beds in the United States. Medical Care, 1997, 35(6):574-588.

(57) Young D. Planning and controlling health capital: attaining an appropriate balance between regulation and competition. Medical Care Review, 1991, 48(3):261-293.

(58) Chohen HS. Regulating Health Care Facilities: the certificate-of-Need process reexamined. Inquiry, 1973, 10:3-9.

(59) Colby DC, Begley CE. The effects of implementation problems on certificate of need decisions in Illinois. Health Policy Education, 1983, April 3(4):303-316.

(60) Consedine JF. Certificate of need and the pitfalls of due process. Inquiry, 1980, Winter 17(4):356.

(61) Mahler J. Barriers to coordinating health services regulatory programs. Journal of Health Politics, Policy and Law, 1981, 6(3):528-541.

(62) Frossard M. Hospital strategy and regional planning in France. International Journal of Health and Management , 1990, 5:59-63.

(63) Hackey RB. New wine in old bottles: certificate of need enters the 1990s. Journal of Health Politics, Policy and Law, 1993, 18(4):927-935.

(64) Scott HD, Tierney JT, Waters WJ, et al. Certificate-of-Need: a state perspective. Rhode Island Medical Journal, 1987, 70(8):341-345.

(65) Shaw J. Hospitals face NY building moratorium. American Medical News, 1983, 26(9):3,26.

(66) Hawryluk M. Problems persist with enforcement rule: surveys find most facilities out of compliance as moratorium ends. Provider, 1995, 21(11):25-26.

(67) Banta D. An approach to the social control of hospital technologies. (Division of Strengthening of Health Services SHS Paper Number 10). Geneva, World Health Organization, 1995. http://whqlibdoc.who.int/hq/1995/WHO_SHS_CC_95.2.pdf

(68) Johansen KS, Racoveanu NT. Big ticket health technology: is rational utilisation possible? Medical Progress Through Technology, 1991, 17(2):85-91.

(69) Fuchs VR. Economics, values and health care reform. American Economic Review, 1996, 86:1-24.

(70) Ramsey SD, Pauly MV. Structural incentives and adoption of medical technologies in HMO and fee-for-service health insurance plans. Inquiry, 1997, 34:228-236.

(71) Banta HD, Luce BR. Health care technology and its assessment: an international perspective. Oxford, Oxford University Press, 1993.

(72) Banta D. Health care technology as a policy issue. Health Policy, 1994, 30:1-21.

(73) Beeson K. Changes in medical therapy during the past half century. Medicine, 1980, 59:79-99.

(74) Hisashige A. Healthcare technology assessment and the challenge to Pharmacoeconomics in Japan. Pharmacoeconomics, 1997, 11(4):319-333.

(75) Maynard A, Bloor K. Help or hindrance? The role of economics in rationing health care. British Medical Bulletin, 1995, 51(4):854-868.

(76) Bloor K, Maynard A, Freemantle N. Lessons from international experience in controlling pharmaceutical expenditure III: regulating industry. British Medical Journal, 1996, 313: 33-35.

(77) Perleth M, Busse R, Shwartz FW. Regulation of health-related technologies in Germany. Health Policy, 1999, 46:105-126.

(78) Lexchin J. Improving the appropriateness of physician prescribing. International Journal of Health Services, 1998, 28(2):253-267.

(79) Banta HD. Medical technology in China. Health Policy, 1990, 14:127-137.

(80) Eastaugh S. Medical economics and health financing. Boston, Auburn House, 1981.

(81) Eastaugh S. Health economics: efficiency, quality and equity. Westport, Auburn House, 1992.

(82) Feldstein M. The rising price of physician services. Review of Economics and Statistics, 1970, 52(2):121-133.

(83) Baumol W. Price controls for medical services and the medical needs of the nation's elderly. Paper commissioned by the American Medical Association, March 11. 1988.

(84) Sloan FA. Physician fee inflation: evidence from the late 1960s. In: Rosett R, ed. The role of health insurance in the health service sector. New York, Watson Academic, 1976: 321-354.

(85) Newhouse J. A model of physician pricing. Southern Economic Journal, 1970, 37(2): 174-183.

(86) Rice T. The impact of changing Medicare reimbursement rates on physician induced demand. Medical Care, 1983, 21:803-815.

(87) Davis K. Economic theories of behaviour in nonprofit, private hospitals. Washington DC, The Brookings Institute, 1971.

(88) Feldstein PJ. Health care economics. New York, John Wiley and Sons, 1979.

(89) Goldfarb M, Hornbrook M, Rafferty J. Behaviour of the multi-product firm: a model of the nonprofit hospital system. Medical Care, 1980, 18(2):185-201.

(90) Hornbrook M, Goldfarb M. A partial test of a hospital behaviour model. Social Science and Medicine, 1983, 17(10):667-680.

(91) Lee ML. A conspicuous production theory of hospital behaviour. Southern Economic Journal, 1971, 38(1):48-58.

(92) Brown M. Economic analysis of hospital operation. Hospital Administration, 1970, 15:60-74.

(93) Klarman H. The Economics of Health. New York, Columbia University Press, 1965.

(94) Long MF. Efficiency use of hospitals. In: Muskin S, ed. The economics of health and medical care. Michigan, Ann Arbor: University of Michigan Press, 1964.

(95) Reder M. Some problems in economics of hospitals. American Economic Review, 1965, 56(2):472-480.

(96) Rice R. Analysis of hospital as an economic organism. The Modern Hospital, 1966, 106(4):88-91.

(97) Clarke R, Gray C. Options for change in the NHS consultant contract. BMJ, 1994, 309: 528-530.

(98) Ron A, Abel-Smith B, Tamburi G. Health insurance in low and middle income countries: the social security approach. Geneva, International Labour Office, 1990.

(99) Normand C, Weber A. Social health insurance: a guidebook for planning. Geneva, World Health Organization, 1994.

(100) Stainwald B. Compensation of hospital-based physicians. Health Service Research, 1983, 18(1):17-47.

(101) Glaser WA. Health insurance in practice: international variations in financing, benefit, and problems. San Francisco, Jossey-Bass Publishers, 1991.

(102) Pontes MC. Agency theory: a framework for analyzing physician services. Health Care Management Review, 1995, 20(4):57-67.

(103) Bunker J. Surgical manpower: a comparison of operations and surgeons in the US, England and Wales. New England Journal of Medicine, 1970, 282:135-144.

(104) McPherson K, Strong PM, Epstein A et al. Regional variations in the use of common surgical procedures. Social Science and Medicine, 1981, 15A:273-288.

(105) Hillman BJ, Joseph CA, Mabry MR et al. Frequency and costs of direct imaging in office practice; a comparison of self-referring and radiologist-referring physicians. New England Journal of Medicine, 1990, 323:1604-1608.

(106) Broomberg J, Price M. The impact of the fee-for-service reimbursement system on the utilization of health services. South Africa Medical Journal, 1990, 78(4):130-132.

(107) Eisenberg J. Physician utilization. Medical Care, 1985, 23:461-483.

(108) Evans RG. Supplier induced demand: some evidence and implications. In: Perlman M, ed. The economics of health and medical care. London, Macmillan, 1974.

(109) Evans RG. Finding the levers, finding the courage: lessons from cost containment in North America. Journal of Health Politics, Policy and Law, 1986, 111:585-615.

(110) Langwell KM, Nelson LM. Physician payment systems: a review of history, alternatives and evidence. Medical Care Review, 1986, 43:5-58.

(111) Greenfield S, Nelson EC, Zubkoff M et al. Variations in resource utilisation among medical specialities and systems of care: results from the medical outcome study. JAMA, 1992, 267(12):1624-1630.

(112) Jacobs P. The Economics of Health and Medical Care. Gaithersburg, An Aspen Publication, 1991.

(113) Donner-Banzhoff N, Kreienbrock L, Katic M et al. Family practitioners' remuneration and patterns of care — does social class matter? Soz Praventivmed, 1998, 43(3):73-79.

(114) Hughes D, Yule B. The effect of per-item fees on the behaviour of general practitioners. Journal of Health Economics, 1992, 11(4):413-437.

(115) Barnum H, Kutzin J, Saxenian H. Incentive and provider payment methods. International Journal of Health Planning and Management, 1995, 10:23-45.

(116) Mills A, Bennett S, Siriwanarangsun P et al. The response of providers to capiation payment: a case-study from Thailand. Health Policy, 2000, 51(3):163-180.

(117) Congressional Budget Office. Effects of managed care: an update. Washington DC, CBO Memorandum, 1994.

(118) Lennon MA, Worthington HV, Coventry P et al. The Capitation Study 2. Does capitation encourage more prevention? British Dental Journal, 1990, 168(5):213-215.

(119) Yesalis CE, Norwood G, Lipson DP et al. Capitation payment for pharmacy services I. Impact on drug use and pharmacist dispensing behaviour. Medical Care, 1984, 22(8): 737-745.

(120) Lipson DP, Yesalis CE, Kohout R et al. Capitation payment for Medicaid pharmacy services: impact on non-Medicaid prescriptions. Medical Care, 1981, 19(3):342-353.

(121) Scott A. Primary or secondary care? What can economics contribute to evaluation at the interface? Journal of Public Health Medicine, 1996, 18(1):19-26.

(122) Schwartz R, et al. Cardiac capitation: vision of the future for speciality care. Cardiology Pre-eminance Roundtable of the Advisory Board Company, 1995.

(123) Culyer AJ, Donaldson C, Gerard K. Financial aspects of health services: drawing on experience. Institute of Health Services Management. Working Paper No 3. York, University of York, 1988.

(124) Rosen B. Professional reimbursement and professional behaviour: emerging issues and research challenges. Social Science and Medicine, 1989, 29(3):455-462.

(125) Gosden T, Pedersen L, Togerson D. How should we pay doctors? A systematic review of salary payments and their effect on doctor behaviour. QJM, 1999, 92(1):47-55.

(126) Donaldson C, Magnussen J. DRGs: the road to hospital efficiency. Health Policy, 1992, 21:47-64.

(127) Rodrigues J. Hospital utilisation and reimbursement method in Brazil. International Journal of Health Planning and Management, 1989, 4:3-15.

(128) Schulenbury JM. Forming and reforming the market for third-party purchasing of health care: a German perspective. Social Science and Medicine, 1993, 39(10):1473-1481.

(129) Robra BP, Swart E, Deh S et al. Rationale of inpatient care — towards a new payment system. Gesundheitswesen, 1999, 61(2):70-77.

(130) Fetter RB. Construction of diagnosis-related groups. Medical Care, 1980, 18(2):5-20.

(131) Lave J. Hospital reimbursement under Medicare. Milbank Memorial Fund Quarterly, 1984, 62(2):251-268.

(132) Lave JR. The effect of the Medicare prospective payment system. Annual Review of Public Health, 1989, 10:141-161.

(133) Omenn GS, Conrad DA. Implication of DRGs for clinicians. New England Journal of Medicine, 1984, 311:1314-1317.

(134) Rosko MD, Broyles RW. Short-term responses of hospitals to the DRG prospective pricing mechanism in New Jersey. Medical Care, 1987, 25:88-94.

(135) Cuterman PB, Dobson A. Impact of the Medicare prospective payment system for hospitals. Health Care Financing Review, 1986, 7:97-114.

(136) Kahn KL, Keeler EB, Sherwood MJ et al. Comparing outcomes of care before and after implementation of the DRG-based prospective payment system. Journal of the American Medical Association, 1990, 264:1984-1988.

(137) Sloan FA, Morrisey MA, Valvona J. Medicare prospective payment and the use of medical technologies in hospitals. Medical Care , 1988, 26:837-853.

(138) Young DW, Saltman RB. Medicare practice, case mix, and cost containment. Journal of the American Medical Association, 1982, 247:801-805.

(139) Chulis GS. Assessing Medicare's prospective payment system for hospitals. Medical Care Review, 1991, 48(2):167-206.

(140) Coulam RF, Gaumer GL. Medicare's prospective payment system: a critical appraisal. Health Care Financing Review, 1992, Annual Supplement:45-77.

(141) Hadley J, Swartz K. The impact on hospital costs between 1980 and 1984 of hospital rate regulation, competition and change in health insurance coverage. Inquiry, 1989, 26:35-47.

(142) Miller M, Sulvetta MB, Englert E. Service mix in the hospital outpatient department: implications for Medicare payment reform. Health Service Research, 1995, 30(1):59-77.

(143) Carrol NV, Erwin WG. Patient shifting as a response to Medicare prospective payment. Medical Care, 1987, 25:1161-1167.

(144) Morrisey MA, Sloan FA, Valvona J. Medicare prospective payment and post-hospital transfers to sub-acute care. Medical Care, 1988, 26:685-698.

(145) Ahern M, McCoy V. Effects of payment changes on urgent/emergent admissions in rural and urban communities. The Journal of Rural Health, 1993, 9(2):120-128.

(146) Blendon RJ, Aiken LH, Freeman HE et al. Uncompensated care by hospitals or public insurance for the poor: does it make a difference? New England Journal of Medicine, 1986, 314:1160.

(147) Wyszewianski L, Thoma W, Friedman B. Case-based payment and the control of quality and efficiency in hospitals. Inquiry, 1987, 24:17-25.

(148) Fitzgerald JF, Fagan LF, Tierny WM et al. Changing patterns of hip fracture care before and after implementation of the prospective payment system. Journal of the American Medical Association, 1987, 258:218-221.

(149) Weinberger M, Ault KA, Vinicor F. Prospective reimbursement and diabetes mellitus: impact on glycaemic control and utilisation of health services. Medical Care, 1988, 26:77-83.

(150) Kahn KL, Rogers W, Rubenstein LV et al. Measuring quality of care with explicit progress criteria before and after implementation of the DRG-based prospective payment system. Journal of the American Medical Association, 1990, 264:1980-1983.

(151) Wiley MM. Hospital financing reform and case mix measurement: an international review. Health Care Financing Review, 1992, 13(4):119-133.

(152) Cromwell J, Dayhoff D, McCall N et al. Medicare participating heart bypass centre demonstration: final report (to the Health Care Financing Administration). Washington, MA, Health Economics Research Inc, 1994.

(153) Duckett S. Hospital payment arrangements to encourage efficiency: the case of Victoria, Australia. Health Policy, 1995, 34:113-134.

(154) Wolfe PR, Moran DW. Global budgeting in the OECD countries. Health Care Financing Review, 1993, 14(3):5576.

(155) Stevens CM. Health care cost containment: some implications of global budgets. Science, 1993, 259:16-17,105.

(156) Hurley JH, Card R. Global physician budget as common-property resources: some implications for physicians and medical associations. Canadian Medical Association Journal, 1996, 154(8):1161-1168.

(157) Hurley J, Lomas J, Goldsmith L. Physician responses to global physician expenditure in Canada: a common property perspective. Milbank Memorial Fund Quarterly, 1997, 75(3):343-363.

(158) Hirdes JP, Botz CA, Kozak J et al. Identifying an appropriate case mix measure for chronic care: evidence from an Ontario pilot study. Healthcare Management Forum, 1996, 9(1):40-46.

(159) Hachey RB. Setting limits through global budget: hospital cost containment in Rhode Island. Spectrum, 1996, 69(1):6-16.

(160) Holahan J, Blumberg LJ, Zuckerman S. Strategies for implementing global budgets. Milbank Memorial Fund Quarterly, 1994, 72(3):399-429.

(161) Bishop CE, Wallack S. National health expenditure limits: the case for a global budget process. Milbank Memorial Fund Quarterly, 1996, 74(3):361-377.

(162) Davies P. Extending the flexible spine of pay. The Health Service Journal, 1988, December:1442.

(163) Buchan J, Thompson M. Pay and nursing performance. Health Manpower Management, 1993, 19(2):29-31.

(164) Buchan J. Performance-related pay and NHS nursing. Nursing Standard, 1993, 7(25):30.

(165) Castledine G. Can performance-related pay be adapted for nursing? British Journal of Nursing, 1993, 2(22):1120-1121.

(166) Scola CM. Keeping your staff. Part 2 - Performance related pay in practice. British Dental Journal, 1990, 169(19):334-336.

(167) Berwich DM. The toxicity of pay for performance. Quality Management in Health Care, 1995, 4(1):27-33.

(168) Bledsoe DR, Leisy WB, Rodeghero JA. Tying physician incentive pay to performance. Healthcare Financial Management, 1995, 49(12):40-44.

(169) Bloor K, Maynard A. Rewarding excellence? Consultants' distinction awards and the need for reform. Centre for Health Economics, Discussion Paper No 100. York, University of York, 1992.

(170) Bloor K, Maynard A, Street A. How much is a doctor worth? Centre for Health Economics, Discussion Paper No 98. York, University of York, 1992.

(171) Griffin RP. Why doesn't performance pay work? Health Manpower Management, 1993, 19(2):11-13.

(172) Hern JE. What should be done about merit award? British Medical Journal, 1994, 308:973-974.

(173) Lewis E. Rewarding performance. The Health Service Journal, 1990, December:1435.

(174) Macara S. Appraisal of doctors in the NHS: merit awards and performance-related pay. British Journal of Hospital Medicine, 1995, 53(3):111-112.

(175) Smith J, Simpson J. Locally determined performance related pay: better levers exist for improving performance in a health service with disparate values. British Medical Journal, 1994, 309:495-496.

(176) Hutchison B. Do physician-payment mechanisms affect hospital utilisation? A study of Health Service Organization in Ontario. Canadian Medical Association Journal, 1996, 154(5):653-661.

(177) Schlackman N. Evolution of a quality-based compensation model: the third generation. American Journal of Medical Quality, 1993, 8(2):103-110.

(178) Hemenway D, Killen A, Cashman SB et al. Physician response to financial incentives. New England Journal of Medicine, 1990, 322:1059-1063.

(179) Winkelman JW, Aitken JL, Wybenga DR. Cost saving in a hospital clinical laboratory with a pay-for-performance incentive program for supervisors. Archives of Pathology and Laboratory Medicine, 1991, 115(1):38-41.

(180) Boyce J, Morris T. Performance related pay for hospital doctors. British Medical Journal, 1992, 305:131-132.

(181) Martin G. Performance-related pay in nursing: theory, practice and prospect. Health Manpower Management, 1994, 20(5):10-17.

(182) Lemieux-Charles L. Physicians in health care management: 2. Management performance: Who, what how and when? Canadian Medical Association Journal, 1994, 150(4):481-485.

(183) Liu X, Mills A. Evaluating payment mechanisms: how can we measure unnecessary care? Health Policy and Planning, 1999, 14(4):409-413.

(184) Rodeghero JA. Benchmarking physicians' practices: trends toward the millennium. Journal of Health Care Financing, 1999, 25(4):15-37.

(185) Chapin R. Incentive payment to nursing homes based on quality-of-care outcomes. Journal of Applied Gerontology, 1992, 11(2):131-145.

(186) Kane R. Outcome-based payment: a new beginning? Health Progress, 1986, 67(3): 36-41, 70.

(187) Langwell K. Price controls: on the one hand ... and on the other. Health Care Financing Review, 1993, 14(3):5-10.

(188) Sloan FA, Hay JW. Medical pricing mechanisms for physician services: an overview of alternative approaches. Medical Care Review, 1986, 43(1):59-100.

(189) McDonough JE. Tracking the demise of state hospital rate setting. Health Affairs, 1997, 16(1):142-149.

(190) Sloan FA. Rate regulation as a strategy for hospital cost control: evidence from the last decade. Milbank Memorial Fund Quarterly, 1983, 61(2):195-217.

(191) Wallack S, Skwara K, Cai J. Redefining rate regulation in a competitive environment. Journal of Health Politics and Law, 1996, 21(3):489-509.

(192) Ikegami N, Ikeda S. The paradox of decreasing prices and increased costs for diagnostic tests, imaging and drugs in Japan. International Journal of Technology Assessment in Health Care , 1997, 13(1):99-110.

(193) Campbell JC, Ikegami N. The art of balance in health policy: maintaining Japan's low-cost, egalitarian system. New York, Cambridge University Press, 1998.

(194) Ikegami N. Japan: maintaining equity through regulated fees. Journal of Health Politics, Policy and Law, 1992, 17(4):689-713.

(195) Oliver A, Ikegami N, Ikeda S. Effect of Japanese government policy on hospital pharmaceutical profit levels. Journal of Health Service Research Policy, 1999, 4(1):27-32.

(196) Saltman RB, Figueras J, Sakellarides C. Critical challenges for health care reform in Europe. Buckingham, Open University Press, 1998.

(197) Jonsson B. Pricing and reimbursement of pharmaceuticals in Sweden. Pharmacoeconomics, 1994, 6(Suppl 1):51-60.

(198) Lopez-Bastida J, Mossialos E. Pharmaceutical expenditure in Spain: cost and control. London, London School of Economics and Social Science, University of London, 1997.

(199) Earl-Slater A. Regulating the price of the UK's drugs: second thoughts after the government's first report. BMJ, 1997, 314:365-368.

(200) Lind J. The reference price system for drugs saves 5 billions crowns in Germany. Lakartidningen, 1992, 89(59):4334-4337.

(201) Giuliani G, Selke G, Garattini L. The German experience in reference pricing. Health Policy, 1998, 44(1):73-85.

(202) Gold M, Chu K, Felt S et al. Effects of selected cost-containment efforts: 1971-1993. Health Care Financing Review, 1993, 14(3):183-213.

(203) Shiell A. Paying for efficiency: what price the quality of hospital care? Australian Journal of Public Health, 1992, 16(3):294-301.

(204) Hsiao WC, Braun P, Yntema D et al. Estimating physicians' work for a resource-based relative-value scale. New England Journal of Medicine, 1988, 319(13):835-841.

(205) Gabel J, Redisch M. Alternative physician payment methods: incentive, efficiency, and national health insurance. Milbank Memorial Fund Quarterly, 1979, 57(1):38-59.

(206) Liu X, Liu Y, Chen N. The Chinese experience of hospital price regulation. Health Policy and Planning, 2000, 15(2):157-163.

(207) Chakraborty S. Lebanon Provider Payment Study. Health Care Financing Administration, 2000.

(208) Wilensky GR. Alleviating the physician glut: what's the government role? JAMA, 1997, 277(1):73.

(209) Armstrong E, Akaho E, Fujii M. Japanese pharmacy: innovation mixed with tradition. The Annals of Pharmacotherapy, 1995, 29:185.

(210) Liu X, Li S. Drug policy in China: transformations, current status and future prospects. Pharmacoeconomics, 1997, 12(1):1-9.

(211) Chen J. The trend of development of modern medicine and the direction of the hospital reform. Chinese Journal of Hospital Management, 1993, 1993(3):11-13.

(212) Dai D. A review to the health care reform. Chinese Journal of Health Economics, 1993, 1993(2):26-28.

(213) Xinhua News Agency. The waste of health resources is serious. People's Daily, 1996, March 7, 1996.

(214) Foster S. Supply and use of essential drugs in sub-Saharan Africa: some issues and possible solutions. Social Science and Medicine, 1991, 32(11):1201-1218.

(215) Le Grand A. Intervention research in rational use of drugs: a review. Health Policy and Planning, 1999, 14(2):89-102.

(216) Trap B, Hansen EH. Cotrimoxazole prescribing by dispensing and non-dispensing doctors: do they differ in rationality? Tropical Medicine and International Health, 2002, 7(10): 878-885.

(217) Trap B, Hansen EH. Treatment of upper respiratory tract infections — a comparative study of dispensing and non-dispensing doctors. Journal of Clinical Pharmacy and Therapeutics, 2002, 27(4):289-298.

(218) Trap B, Hansen EH, Hogerzeil HV. Prescription habits of dispensing and non-dispensing doctors in Zimbabwe. Health Policy Plan 2002, 2002, 17(3):288-295.

(219) Brown ER. Should single-payer advocates support President Clinton's proposal for health care reform? American Journal of Public Health, 1994, 84(2):182-185.

(220) Woolhandler S, Himmelstein D. The deteriorating administrative efficiency of the US health care system. New England Journal of Medicine, 1991, 324:1253-1258.

(221) McDermott J. Evaluating health system reform: The case for a single-payer approach. Journal of American Medical Association, 1994, 271(10):782-784.

(222) Culyer AJ, Meads A. The United Kingdom: effective, efficient, equitable? Journal of Health Politics, Policy and Law, 1992, 17(4):667-687.

(223) Evans RG. Canada: the real issues. Journal of Health Politics, Policy and Law, 1992, 17(4):739-761.

(224) Busse R, Schwartz FW. Financing reforms in the German hospital sector: from full cost cover principle to prespective case fees. Medical Care, 1997, 35(19) supplement:OS40-49.

(225) Weil TP. Comparison of medical technology in Canadian, German and US hospitals. Hospital and Health Services Administration, 1995, 40(4):524-533.

(226) Ayah RT. Impact of user fees in health. East Africa Medical Journal, 1997, 74:749-750.

(227) de Ferranti D. Paying for health services in low and middle income countries. World Bank Working Paper No 721. Washington DC, World Bank, 1985.

(228) Akin J, Birdsall N, de Ferranti D. Financing health services in low and middle income countries: an agenda for reform. Washington DC, World Bank Publications, 1987.

(229) Litvack JI, Bodart C. User fees plus quality equals improved access to health care: results of a field experiment in Cameroon. Social Science and Medicine, 1993, 37(3):383.

(230) Gertler P, Locay L, Sanderson W. Are user fees regressive? Journal of Econometrics, 1987, 36:67-88.

(231) Waddington C, Enyimayew W. A price to pay: the impact of user charges in Ashanto-Akim district, Ghana. International Journal of Health Planning and Management, 1989, 4:17-48.

(232) Yoder R. Are people willing and able to pay for health services? Social Science and Medicine, 1989, 29:35-42.

(233) De Bethune X, Alfani S, Lahaye J. The influence of an abrupt price increase on health service utilisation: evidence from Zaire. Health Policy and Planning, 1989, 4:76-81.

(234) Hussein AK, Mujinja PG. Impact of user charges on government health facilities in Tanzania. East Africa Medical Journal, 1997, 74(12):751-757.

(235) Huber JH. Ensuring access to health care with the introduction of user fees: a Kenyan example. Social Science and Medicine, 1993, 36(4):485-494.

(236) Chabot J. The Bamako Initiative. Lancet, 1988, 8624:1366-1367.

(237) Creese AL. User charges for health care: a review of recent experience. Health Policy and Planning, 1991, 6:309-319.

(238) Kanji N. Charging for drugs in Africa: UNICEF's Bamako Initiative. Health Policy and Planning, 1989, 4(2):110-120.

(239) Korte R, Richter H, Merkle F et al. Financing health services in sub-Saharan Africa: options for decision makers during adjustment. Social Science and Medicine, 1992, 29:1-9.

(240) Owa JA, Osinaiki AI, Costello AM. Charging for health services in low and middle income countries. Lancet, 1992, 340:732.

(241) Stanton B, Clemens J. User fees for health care in low and middle income countries: a case study of Bangladesh. Social Science and Medicine, 1989, 10:1199-1205.

(242) Fielder JL, Suazo J. Ministry of Health user fees, equity and decentralization: lessons from Honduras. Health Policy Plan 2002, 2002, December 17(4):362-377.

(243) Knippenberg R, Levy-Bruhl D, Osseni R et al. Children in the tropics: the Bamako Initiative primary health care experience. Paris, International Children's Centre, 1990.

(244) Russell S, Gilson L. User fee policies to promote health service access for the poor: a wolf in sheep's clothing. International Journal of Health Services, 1997, 27(2):359-379.

(245) Gilson L, Russell S, Buse K. The political economy of user fees with targeting: developing equitable health financing policy. Journal of International Development, 1995, 7(3): 369-401.

(246) Chisadza E, Maponga CC, Nazerali H. User fees and drug pricing policies: a study at Harare Central Hospital, Zimbabwe. Health Policy and Planning, 1995, 10(3):319-326.

(247) Gilson L. The lessons of user fee experience in Africa. Health Policy and Planning, 1997, 12(4):273-285.

(248) Kutzin J. The appropriate role for patient cost sharing. Buckingham, Open University Press, 1998.

(249) Rubin RI, Mandelson DN. A framework for cost-sharing policy analysis. In: Mattison N, ed. Sharing the costs of health: a multicountry perspective. Basle, Pharmaceutical Partner for Better Health Care, 1995.

(250) Rice T, Morrison KR. Patient cost sharing for medical services: a review of the literature and implications for health care reform. Medical Care Review, 1994, 51(3):235-287.

(251) Abel-Smith B, Figuras J, Holland A. Choice in health policy: an agenda for the European Union. Luxembourg and Aldershot: Office for Official Publications of the European Communities and Dartmouth Publishing Company, 1995.

(252) Walley T, Haycox A, Barton S. Drug rationing in the UK National Health Service: current status and future prospects. Pharmacoeconomics, 1997, 12(3):339-350.

(253) Scitovsky AA, Snyder NM. Effect of coinsurance on use of physician services. Social Security Bulletin, 1972, 35(6):3-19.

(254) Scitovsky AA, McCall. Coinsurance and the demand for physician services: four years later. Social Security Bulletin, 1977, 40:19-27.

(255) Manning WG, Newhouse J, Duan N et al. Health insurance and the demand for medical care: Evidence from a randomized experiment. American Economic Review, 1987, 77: 251-277.

(256) Scheffler RM. United Mine Workers health plan: an analysis of the cost sharing program. Medical Care, 1984, 22(3):247-254.

(257) Cherkin DG, Grothaus L, Wagner H. The effect of office visit copayments on utilization in a Health Maintenance Organization. Medical Care, 1989, 27(7):24-38.

(258) Cherkin DG, Grothaus L, Wagner H. The effect of office visit copayments on preventive care services in an HMO. Inquiry, 1990, 27(1):24-38.

(259) Selby JV, Fireman BH, Swain BE. Effect of a copayment on use of the emergency department in a health maintenance organization. The New England Journal of Medicine, 1996, 334(10):635-641.

(260) Stuart B, Zacker C. Who bear the burden of Medicaid drug copayment policies. Health Affairs, 1999, 18(2):201-213.

(261) Harris BL, Stergachis A, Ried LD. The effect of drug copayment on utilization and cost of pharmaceuticals in a health maintenance organization. Medical Care, 1990, 28:907-917.

(262) Bech RG, Horne JM. Utilization of public insured health services in Saskatchewan before, during and after copayment. Medical Care, 1980, 18:787-806.

(263) Kupor SA, Liu YC, Lee J et al. The effect of copayments and income on the utilisation of medical care by subscriber to Japan's national health insurance system. International Journal of Health Services, 1995, 25:295-312.

(264) Nolan B. Economic incentives, health status and health services utilisation. Journal of Health Economics, 1993, 12:151-169.

(265) OECD. The reform of health care systems: a review of seventeen OECD countries. Paris, Organisation for Economic Co-operation and Development, 1994.

(266) Ryan M, Birch S. Charging for Health care: evidence on the utilisation of NHS prescribed drugs. Social Science and Medicine, 1991, 33:681-687.

(267) Lohr KN, Brook RH, Kamberg CJ et al. Use of medical care in the RAND Health Insurance Experiment: diagnosis-and-service-specific analyses in a randomised controlled trial. Medical Care, 1986, 24(Supplement):1-87.

(268) Brian EW, Gibbens SF. California's Medi-Cal copayment experiment. Medical Care, 1974, 12:4-56.

(269) Roemer MI, Hopkins CE, Carr L et al. Copayments for ambulatory care: penny-wise and pound-foolish. Medical Care, 1975, 13(6):457-466.

(270) Freemantle N, Bloor K. Lessons for international experience in controlling pharmaceutical expenditure. I: influencing patients. British Medical Journal, 1996, 312:1469-1471.

(271) Welch BL, Hay JW, Miller DS, et al. The RAND Health Insurance Study: a summary of critique. Medical Care, 987, 25(2):148-156.

(272) Hsiao WC. Primer on health care policy for macroeconomists (prepared for Fiscal Affairs Department, IMF). Boston MA, Harvard University School of Public Health, 1999.

(273) Fahs MC. Physician response to the United Mine Workers' cost-sharing program: the other side of the coin. Health Service Research, 1992, 27(1):25-45.

(274) Jonsson B, Gerdtham UG. Cost sharing for pharmaceuticals: the Swedish reimbursement system. In: Mattison N, ed. Sharing the costs of health: a multicountry perspective. Basle, Pharmaceutical Partner for Better Health Care, 1995.

(275) US Congress O. Does health insurance make a difference? (background paper OTA-BO H-99). Washington DC, US Government Printing Office, 1992. http://govinfo.library.unt.edu/ota/Ota_1/DATA/1992/9212.PDF

(276) Rasell ME. Cost sharing in health insurance: a re-examination. New England Journal of Medicine, 1995, 332:1164-1168.

(277) Maxwell RJ. Why rationing is on the agenda. British Medical Bulletin, 1995, 51(4): 761-768.

(278) Klein R, Day P, Redmayne S. Ration in the NHS: the dance of the seven veils — in reverse. British Medical Bulletin, 1995, 51(4):796-800.

(279) Kitzhaber J, Kemmy A. On the Oregon trail. British Medical Bulletin, 1995, 51(4):818.

(280) Jacobs L, Marmor T, Oberlander J. The Oregon health plan and the political paradox of rationing: what advocates and critics have claimed and what Oregon did. Journal of Health Politics, Policy and Law, 1999, 24(1):161-180.

(281) Tengs TO, Meyer G, Siegel JE et al. Oregon's Medicaid ranking and cost-effectiveness: is there any relationship? Medical Decision Making, 1996, 16(2):99-107.

(282) Burda D. Oregan estimates rationing will save 22%. Modern Health Care, 1991, May 6:12.

(283) Hadorn D. The Oregon priority-setting exercise: cost-effectiveness and the rule of rescue, revisited. Medical Decision Making, 1996, 16(2):117-119.

(284) Leichter HM. Oregon's bold experiment: whatever happened to rationing. Journal of Health Politics, Policy and Law, 1999, 24(1):147-160.

(285) Harrison S. A policy agenda for health care rationing. British Medical Bulletin, 1995, 51(4):885-899.

(286) Bochner F, Martin ED, Burgess NG et al. How can hospitals ration drugs? Drug rationing in a teaching hospital: a method to assign priorities. BMJ, 1994, 308:901-905.

(287) Donaldson C. How can hospital ration drugs? Formulate, don't formularise. BMJ, 1994, 308:905-906.

(288) van de Ven M. Choices in health care: a contribution from the Netherlands. British Medical Bulletin, 1995, 51(4):781-790.

(289) Cooper MH. Core services and the New Zealand health reforms. British Medical Bulletin, 1995, 52(4):799-807.

(290) Calltorp J. Sweden: no easy choices. British Medical Bulletin, 1995, 52(4):791-798.

(291) Conviser R, Retondo MJ, Loveless MO. Predicting the effect of the Oregan Health Plan on Medicaid coverage for outpatient with HIV. American Journal of Public Health, 1994, 84(2):1994-1996.

(292) Ham C. Synthesis: what can we learn from international experience? British Medical Bulletin, 1995, 51(4):819-830.

(293) Musgrove P. Public spending on health care: how are different criteria related. Health Policy, 1999, 47:207-223.

(294) Tarimo E. Essential health service packages: uses, abuse and future directions. Geneva, World Health Organization (Division of Analysis, Research and Assessment), 1997.

(295) Callahan D. What is a reasonable demand on health care resources? Designing a basic package of benefits. Journal of Contemporary Health Law and Policy, 1992, 8(1):1-12.

(296) Bobadilla JL, Cowley P, Musgrove P et al. Design, content and financing of an essential national package of health services. Bulletin of the World Health Organization, 1994, 72(4):653-662.

(297) World Health Organization. The use of essential drugs: eighth report of the WHO Expert Committee. (WHO Technical Report Series No 882). Geneva, World Health Organization, 1998. http://whqlibdoc.who.int/trs/WHO_TRS_882.pdf

(298) Adikwu M, Osondu B. Four years of essential drugs' list in Nigeria. Social Science and Medicine, 1991, 33(9):1005-1010.

(299) Saunders P, Herxheimer A. WHO's model list of essential drugs: a beacon through two decades. Lancet, 1995, 346:519.

(300) Howard N, Laing R. Changes in the World Health Organization essential drug list. Lancet, 1991, 338:743-745.

(301) World Health Organization. Health System Profile Data Base. Geneva, World Health Organization, 2000.

(302) McPake B, Hanson K, Mills A. Community financing of health care in Africa: an evaluation of the Bamako Initiative. Social Science and Medicine, 1993, 36(11):1383-1395.

(303) Ceesay N, Kalliecharan RV. Sustaining the Bamako Initiative in the Gambia: some management implications. World Hospitals and Health Services, 1994, 30(3):10-14.

(304) Krause G, Benzler J, Heinmüller R et al. Performance of village pharmacies and patient compliance after implementation of an essential drug programme in rural Burkina Faso. Health Policy and Planning, 1998, 13(2):159-166.

(305) Levy-Bruhl D, Soucat A, Osseni R et al. The Bamako Initiative in Benin and Guinea: improving the effectiveness of primary health care. International Journal of Health Planning and Management, 1997, 12(Supplement 1):S49-79.

(306) Ogunbekun I, Adeyi O, Wouters A et al. Cost and financing of improvements in the quality of maternal health services through the Bamako Initiative I Nigeria. Health Policy and Planning, 1996, 11(4):369-384.

(307) Tran TT, Nguyen TN, Le T, et al. Establishment of drug chests in commune health station in Vietnam, Bamako Initiative. Southeast Asian Journal of Medicine and Public Health, 1998, 29(3):628-635.

(308) Hogerzeil HV, Walker GJA, Sallami AO et al. Impact of an essential drugs programme on availability and rational use of drugs. Lancet, 1989, 1(8630):141-142.

(309) Woolf S, Grol R, Hutchinson A et al. Potential benefit, limitations, and harms of clinical guidelines. British Medical Journal, 1999, 318:527-530.

(310) Wennberg JE, Barnes BA, Zubkoff M. Professional uncertainty and the problems of supplier-induced demand. Social Science and Medicine, 1982, 16:811-824.

(311) Foreman J. Practice parameters designed to reduce unnecessary medical care. Archives of Ophthalmology, 1990, June:789.

(312) Larson EB. Evidence-based medicine: is translating evidence into practice a solution to the cost-quality challenges facing medicine? The Joint Commission Journal of Quality Improvement, 1999, 25(9):480-485.

(313) Hayes OW. Clinical practice guidelines: a review. Journal of American Osteopath Association, 1994, 94(9):732-738.

(314) Bloor K, Freemantle N. Lessons from international experience in controlling pharmaceutical expenditure II: influencing doctors. British Medical Journal, 1996, 312:1525-1527.

(315) Bray KK, Williams K. An evidence-based approach to clinical practice guidelines in periodontics. Current Opinion in Peridontology, 1997, 4:96-103.

(316) Pestotnik SL, Classen DC, Evans RS et al. Implementing antibiotic practice guidelines through computer-assisted decision support: clinical and financial outcomes. Annals of Internal Medicine, 1996, 124(19):884-890.

(317) Grimshaw JM, Russell IT. Effect of clinical guidelines on medical practice: a systematic review of rigorous evaluations. Lancet, 1993, 342:1317-1322.

(318) Citrome L. Practice protocols, parameters, pathways, and guidelines: a review. Administration and Policy in Mental Health, 1998, 25(3):257-269.

(319) Olmstead T, Zeckhauser R. The menu-setting problem and subsidised prices: drug formulary illusion. Journal of Health Economics, 1999, 18:523-550.

(320) Grabowski H, Mullins C. Pharmacy benefit management, cost-effectiveness analysis and drug formulary decisions. Social Science and Medicine, 1997, 45(4):535-544.

(321) Bos M. Health care technology in the Netherlands. Health Policy, 1994, 30:207-255.

(322) Rothberg A, Walters L. Formulary and funding implications of the gap between the national essential drugs list and current prescribing in a large health maintenance organisation. South Africa Medical Journal, 1996, 86(9):1084-1090.

(323) Chang S, Wong JW, Wong CW, et al. Overview of the hospital formulary systems in Hong Kong. The Annals of Pharmacotherapy, 1997, 31:1526-1530.

(324) Hazlet TK, Hu T. Association between formulary strategies and hospital drug expenditures. American Journal of Hospital Pharmacy, 1992, 49:2207-2210.

(325) Gross DJ. Prescription drug formularies in managed care: concerns for the elderly population. Clinical Therapy, 1998, 20(6):1277-1291.

(326) Hux JE, Naylor CD. Drug prices and third party payment: do they influence medication selection? Pharmacoeconomics, 1994, 5(4):343-350.

(327) Brody B, Wray N, Bame S et al. The impact of economic considerations on clinical decision making: the case of thrombolytic therapy. Medical Care, 1991, 29:899-910.

(328) Kozma CM, Reeder CE, Lingle EW. Expanding Medicaid drug formulary coverage: effects on utilisation of related service. Medical Care, 1990, 28:963-977.

(329) Safavi KT, Hayward RA. Choosing between apples and apples: physicians' choices of prescription drugs that have similar side effects and efficacy. Journal of General Internal Medicine, 1992, 7:32-37.

(330) Soumerai SB, Avorn J, Ross-Degnan D et al. Payment restrictions for prescription drugs under Medicaid: effects on therapy, cost, and equity. New England Journal of Medicine, 1987, 317:550-556.

(331) Monane M, Nagle B, Kelly MA. Pharmacotherapy: strategies to contraol drug costs in managed care. Geriatrics, 1998, 53(9):52-63.

(332) Horn S, Sharkey PD, Phillips-Harris C. Formulary limitations in the elderly: results from the managed care outcomes project. American Journal of Managed Care, 1998, 468:12-29.

(333) Horn S. Unintended consequences of drug formulary. American Journal of Health System and Pharmacy, 1996, 53:2204-2206.

(334) Zusman J. Utilisation review: theory, practice, and issues. Hosptal and Community Psychiatry, 1990, 41(5):531-536.

(335) Mooney G, Ryan M. Rethinking medical audit: the goal is efficiency. Journal of Epidemiology and Community Health, 1992, 46:180-183.

(336) Hopkins A. Approaches to medical audit. Journal of Epidemiology and Community Health, 1991, 45:1-3.

(337) Restuccia JD. The evolution of hospital utilisation review methods in the United States. Social Science and Medicine, 1995, 7(3):253-260.

(338) Butman SM. Precertification for percutaneous transluminal coronary angioplasty in Medicare beneficiaries: a melting pot or a need for better national standards? Catheterization and Cardiovascular Diagnosis, 1990, 21(4):227-232.

(339) Gertman PM, Restuccia JD. The appropriateness evaluation protocol: a technique for assessing unnecessary days of hospital care. Medical Care, 1981, 19(8):855-871.

(340) Wolff N, Schlesinger M. Risk, motives, and styles of utilisation review: a cross-condition comparison. Social Science and Medicine, 1998, 47(7):911-926.

(341) Lorenzo S, Beech R, Lang T et al. An experience of initialisation review in Europe: sequel to a BIOMED project. International Journal for Quality in Health Care, 1999, 11(1):13-19.

(342) Kirkman B, Van de Van W. Improving efficiency in the Dutch health care system: current innovations and future options. Health Policy, 1989, 13:35-53.

(343) Wichizer T. The effect of utilisation review on hospital use and expenditures: a review of the literature and an update on recent findings. Medical Care Review, 1990, 47(3): 327-363.

(344) Dubois RW. Reducing unnecessary care: different approaches to the 'big ticket' and the 'little ticket' items. Journal of Ambulatory Care Management, 1991, 14(4):30-37.

(345) Hayes RP, Lundberg MT, Ballard DJ. Peer review organizations: scientific challenges in HCFA's health care quality improvement initiative. Medical Care Review, 1994, 51(1):39-60.

(346) Groves RE. Therapeutic drug-use review for the Florida Medicaid Program. American Journal of Hospital Pharmacy, 1985, 42:316-319.

(347) Huntsinger L. Drug use review of Oregon. Pharmacy Times, 1990, May:39-41.

(348) Lipowski EE, Collins TM. Wisconsin Medicaid DUR project. The Wisconsin Pharmacist, 1988, 57:20,30,37.

(349) Nycamp D. Impact of drug-regimen review on the Georgia Medicaid Program. The Consultant Pharmacist, 1990, 5(4):235-238.

(350) Peters J. Oregon Medicaid drug use review. Oregon Pharmacist, 1988, 36:36.

(351) Pollard MR, Coster JM. Saving for Medicaid drug spending. Health Affairs, 1991, 10: 196-206.

(352) The Medi-Cal Therapeutic Drug Utilization Review Project: Quarterly Report and Final Analysis Report. Menlo Park, CA, SRI International, 1991.

(353) Moore W. Medicaid drug utilisation review: a critical appraisal. Medical Care Review, 1994, 51(1):3-37.

(354) Bochner F, Burgess NG, Martin ED. Approaches to rationing drugs in hospitals. Pharmacoeconomics, 1996, 10(5):467-474.

(355) Greco PJ, Eisenberg JM. Changing physicians' practices. New England Journal of Medicine, 1993, 329:1271-1274.

(356) Blackburn JL. Impact of drug usage review on drug utilisation. Pharmacoeconomics, 1993, 3(1):14-21.

(357) O'Connell DL, Henry D, Tomlins R. Randomised controlled trial of effect of feedback on general practitioners' prescribing in Australia. British Medical Journal, 1999, 318:507-511.

(358) Schaffner W, Ray WA, Federspiel CF et al. Improving antibiotic prescribing in office practice: a controlled trial of three education method. JAMA, 1983, 250:1728-1732.

(359) Soumerai SB, Lexchin J. Economic and policy analysis of university-based drug detailing. Medical Care, 1986, 24:313-331.

(360) Brand D, Quam L, Leatherman S. Medical practice profiling: concepts and caveats. Medical Care Research and Review, 1995, 52(2):223-251.

(361) Doing what counts for patient safety: federal actions to reduce medical errors and their impact. Washington, Quality Interagency Coordination Task Force, 2000.

(362) Institute of Medicine. To err is human: building a safer health system. Washington DC, National Academy Press, 1999.

(363) Brennan TA, Leape LL, Laird NM. Incidence of adverse events and negligence in hospitalized patients: results of the Harvard Medical Practice Study I. New England Journal of Medicine, 1991, 324:370-376.

(364) Thomas EJ, Studdert DM, Newhouse J. Costs of medical injuries in Utah and Colorado. Inquiry, 1999, 36:255-264.

(365) Leape LL. Error in medicine. Journal of American Medical Association, 1994, 272:1851-1857.

(366) Bates DW, Cullen DJ, Laird N et al. Incidence of adverse drug events and potential adverse drug events: implications for prevention. Journal of American Medical Association, 1995, 274(1):29-34.

(367) Leape LL, Brennan TA, Laird N et al. The nature of adverse events in hospitalized patients: results of the Harvard Medical Practice Study II. New England Journal of Medicine, 1991, 324:377-384.

(368) Bradley EH, Horwitz SM, Grogan CM et al. Monitoring clinical quality in Medicaid managed care. Connecticut Medicine, 1998, 62(4):215-220.

(369) Gold M, Felt S. Reconciling practice and theory: challenges in monitoring Medicaid managed care quality. Health Care Financing Review, 1995, 16(4):85-105.

(370) Brook RH, Lohr KN. Monitoring quality of care in the Medicare program: two proposed systems. Journal of American Medical Association, 1987, 258(21):3138-3141.

(371) Joint Commission on Accreditation of Healthcare Organisations. Agenda for change update, Volume I, No I. Chicago, JCAHO, 1987.

(372) Miller RH, Luft H. Managed care plan performance since 1980. JAMA, 1994, 271(19): 1512-1519.

(373) Lyles A, Palumbo FB. The effect of managed care on prescription drug costs and benefits. Pharmacoeconomics, 1999, 15(2):129-140.

374) Steiner A, Robinson R. Managed care: US research evidence and its lessons for the NHS. Journal of Health Service Research Policy, 1998, 3(3):173-184.

(375) Thabrany H. Prepaid health care in Indonesia. Paper presented at the Second World Conference of the International Health Economics Association, Rotterdam, 6-9 June. 1999. Ref Type: Report

(376) Zweifel P. Managed care in Germany and Switzerland: two approaches to a common problem. Pharmacoeconomics, 1998, 14(Suppl 1):1-8.

(377) Zwanziger J, Melnick GA. Can managed care plans control health care costs? Health Affairs, 1996, 15(2):185-199.

(378) Manning WG, Leibowitz A, Goldberg GA et al. A controlled trial of the effect of a prepaid group practice on use of services. New England Journal of Medicine, 1984, 310:1505-1510.

(379) Mauldon J, Leibowitz A, Buchanan JL et al. Rationing or rationalizing children's medical care: comparison of a Medicaid HMO with fee-for-service care. American Journal of Public Health, 1994, 84(6):899-904.

(380) Miller RH, Luft H. Does managed care lead to better or worse quality of care? Health Affairs, 1997, 16(5):7-25.

(381) Miller EA, Weissert WG, Chernew M. Managed care for elderly people: a compendium of findings. American Journal of Medical Quality, 1998, 13(3):127-140.

(382) Rossiter LF, Nelson LM, Adamache KW. Service use and costs for Medicare beneficiaries in risk-based HMOs and CMPs: some interim results from the National Medicare Competition Evaluation. American Journal of Public Health, 1988, 78:937-943.

(383) Garfinkel SA. Choice of payment plan in the Medicare Capitation Demonstration. Medical Care, 1986, 24:628-640.

(384) Brown RS, Clement DG, Hill JW, et al. Do health maintenance organisations work for Medicare? Health Care Financing Review, 1993, 15:7-23.

(385) McCall N. Evaluation of the Arizona Health Care Cost Containment System Demonstration: First Outcome Report. San Francisco, Laguna Research Associate, 1991.

(386) Wrightson CW, McCall N. Evaluation of Arizona's health care cost containment system: lessons from an eleven-year evaluation. Advanced Health Economics and Health Service Research, 1995, 15:79-97.

(387) Retchin SM, Brown B. Management of colorectal cancer in Medicare health maintenance organization. Journal of General Internal Medicine, 1990, 5:110-114.

(388) Hellinger F. The effect of managed care on quality. Archive of Internal Medicine, 1998, 158:833-841.

(389) Riley GF, Potosky AL, Lubitz JD et al. Stage of cancer diagnosis for Medicare HMO and fee-for-service enrolees. American Journal of Public Health, 1994, 84:1598-1604.

(390) Carlisle DM, Siu AL, Keeler EB et al. HMO vs FFS care of older persons with acute myocardial infarction. American Journal of Public Health, 1992, 82:1626-1630.

(391) Clement DG, Retchin SM, Brown RS et al. Access and outcomes of elderly patients enrolled in managed care. Journal of American Medical Association, 1994, 271: 1487-1492.

(392) Manton KG, Newcomer R, Lowrimore GR et al. Social/health maintenance organisation and fee-for-service health outcomes over time. Health Care Financing Review, 1993, 15:173-202.

(393) Ware JE, Bayliss MS, Rogers WH et al. Differences in 4-year health outcomes for elderly patients enrolled in managed care. Journal of American Medical Association, 1996, 276:1039-1047.

(394) Shaughnessy PW, Schlenker RE, Hittle DF. Home health care outcomes under capitated and fee-for-service payment. Health Care Financing Review, 1994, 16:187-221.

(395) Rogers W, Wells K, Meredith L. Outcome for adult outpatients with depression under prepaid or fee-for-service financing. Archives of General Psychiatry, 1993, 50(7):517-525.

(396) Shaughnessy PW, Schlenker RE, Hittle DF. Case mix of home health patients under capitated and fee-for-service payment. Health Service Research, 1995, 30(1):79-113.

(397) Giffin R. Can managed care save Medicare? Achieving high quality and low costs through managed care. Managed Care Quarterly, 1996, 4(4):12-29.

(398) Dembe A. Preserving workers' compensation benefits in a managed health care environment. Journal of Public Health Policy, 1998, 19(2):200-218.

(399) Murray CJL, Lopez AD. The Global Burden of Disease. Harvard School of Public Health, 1996.

(400) Murray CJL, Lopez AD, Mathers CD et al. The Global Burden of Disease 2000 project: aims, methods and data sources. GPE Discussion Paper No 36. Geneva, World Health Organization, 2000. http://www3.who.int/whosis/discussion_papers/pdf/paper36.pdf

(401) Wennberg JE. The paradox of appropriate care. Journal of American Medical Association, 1987, 258:2568-2575.

(402) Navarro V. (1994) The need to mobilise support for the Wellstone-McDermott Conyer single-payer proposal. American Journal of Public Health 84(2):178-9

(403) Terris M (1998) Lean and mean: the quality of care in the era of managed care. Journal of Pubic Health policy 19(1):5-14.

Index Tables and illustrations are indicated by bold page numbers.