

## CHAPTER ONE

# *Why do Health Systems Matter?*

*Health systems consist of all the people and actions whose primary purpose is to improve health. They may be integrated and centrally directed, but often they are not. After centuries as small-scale, largely private or charitable, mostly ineffectual entities, they have grown explosively in this century as knowledge has been gained and applied. They have contributed enormously to better health, but their contribution could be greater still, especially for the poor. Failure to achieve that potential is due more to systemic failings than to technical limitations. It is therefore urgent to assess current performance and to judge how health systems can reach their potential.*

---

# 1

## WHY DO HEALTH SYSTEMS MATTER?

### THE CHANGING LANDSCAPE

On 13 October 1999, in a maternity clinic in Sarajevo, Helac Fatima gave birth to a son. This was a special occasion, because United Nations demographers had calculated the global population would reach six billion on that day. The little Sarajevo boy was designated as the sixth billionth person on the planet.

Today there are four times as many people in the world as there were 100 years ago – there are now about 4000 babies born every minute of every day – and among the countless, bewildering changes that have occurred since then, some of the most profound have occurred in human health. For example, few if any of Helac Fatima’s ancestors around 1899 were likely to have seen a hospital, far less been born in one.

The same was true for the great majority of the 1.5 billion people then alive. Throughout the world, childbirth invariably occurred at home, rarely with a physician present. Most people relied on traditional remedies and treatments, some of them thousands of years old. Most babies were born into large families and faced an infancy and childhood threatened by a host of potentially fatal diseases – measles, smallpox, malaria and poliomyelitis among them. Infant and child mortality rates were very high, as were maternal mortality rates. Life expectancy for adults was short – even half a century ago it was a mere 48 years at birth.

Last year the son of Helac Fatima entered the world with a life expectancy at birth of 73 years – the current Bosnian average. The global average is 66 years. He was born in a big city hospital staffed by well-trained midwives, nurses, doctors and technicians – who were supported by modern equipment, drugs and medicines. The hospital is part of a sophisticated health service. It is connected in turn to a wide network of people and actions that in one way or another are concerned with maintaining and improving his health for the rest of his life – as for the rest of the population. Together, all these interested parties, whether they provide services, finance them or set policies to administer them, make up a health system.

Health systems have played a part in the dramatic rise in life expectancy that occurred during the 20th century. They have contributed enormously to better health and influenced the lives and well-being of billions of men, women and children around the world. Their role has become increasingly important.

Enormous gaps remain, however, between the potential of health systems and their actual performance, and there is far too much variation in outcomes among countries which seem to have the same resources and possibilities. Why should this be so? Health systems would seem no different from other social systems in facing demands and incentives to

perform as well as possible, and it might be expected that – with some degree of regulation by the state – their performance could be largely left to markets, just as with the provision of most other goods and services.

But health is fundamentally different from other things that people want, and the difference is rooted in biology. As eloquently expressed by Jonathan Miller, “Of all the objects in the world, the human body has a peculiar status: it is not only possessed by the person who has it, it also possesses and constitutes him. Our body is quite different from all the other things we claim as our own. We can lose money, books and even houses and still remain recognisably ourselves, but it is hard to give any intelligible sense to the idea of a disembodied person. Although we speak of our bodies as premises that we live in, it is a special form of tenancy: our body is where we can always be contacted” (1). The person who seeks health care is of course a consumer – as with all other products and services – and may also be a co-producer of his or her health, in following good habits of diet, hygiene and exercise, and complying with medication or other recommendations of providers. But he or she is also the physical object to which all such care is directed.

Health, then, is a characteristic of an inalienable asset, and in this respect it somewhat resembles other forms of human capital, such as education, professional knowledge or athletic skills. But it still differs from them in crucial respects. It is subject to large and unpredictable risks, which are mostly independent of one another. And it cannot be accumulated as knowledge and skills can. These features are enough to make health radically unlike all other assets which people insure against loss or damage, and are the reason why health insurance is more complex than any other kind of insurance. If a car worth US\$ 10 000 would cost \$15 000 to repair after an accident, an insurer would only pay \$10 000. The impossibility of replacing the body, and the consequent absence of a market value for it, precludes any such ceiling on health costs.

Since the poor are condemned to live in their bodies just as the rich are, they need protection against health risks fully as much. In contrast, where other assets such as housing are concerned, the need for such protection either does not arise, or arises only in proportion to income. This basic biological difference between health and other assets even exaggerates forms of market failure, such as moral hazard and imperfect and asymmetric information, that occur for other goods and services. Directly or indirectly, it explains much of the reason why markets work less well for health than for other things, why there is need for a more active and also more complicated role for the state, and in general why good performance cannot be taken for granted.

The physical integrity and dignity of the individual are recognized in international law, yet there have been shameful instances of the perversion of medical knowledge and skills, such as involuntary or uninformed participation in experiments, forced sterilization, or violent expropriation of organs. Health systems therefore have an additional responsibility to ensure that people are treated with respect, in accordance with human rights.

This report sets out to analyse the role of health systems and suggest how to make them more efficient and, most importantly, more accessible and responsive to the hundreds of millions of people presently excluded from benefiting fully from them. The denial of access to basic health care is fundamentally linked to poverty – the greatest blight on humanity’s landscape. For all their achievements and good intentions, health systems have failed globally to narrow the health divide between rich and poor in the last 100 years. In fact, the gap is actually widening. Some such worsening often accompanies economic progress, as the already better-off are the first to benefit from it. But the means exist to accelerate the sharing by the poor in these benefits, and often at relatively low cost (see Box 1.1). Finding

a successful new direction for health systems is therefore a powerful weapon in the fight against poverty to which WHO is dedicated. Not least for the children of the new century, countries need systems that protect all their citizens against both the health risks and the financial risks of illness.

## WHAT IS A HEALTH SYSTEM?

In today's complex world, it can be difficult to say exactly what a health system is, what it consists of, and where it begins and ends. This report defines a health system to include *all the activities whose primary purpose is to promote, restore or maintain health.*

Formal health services, including the professional delivery of personal medical attention, are clearly within these boundaries. So are actions by traditional healers, and all use of medication, whether prescribed by a provider or not. So is home care of the sick, which is how somewhere between 70% and 90% of all sickness is managed (2). Such traditional public health activities as health promotion and disease prevention, and other health-enhancing interventions like road and environmental safety improvement, are also part of the system. Beyond the boundaries of this definition are those activities whose primary purpose is something other than health – education, for example – even if these activities have a secondary, health-enhancing benefit. Hence, the general education system is outside the boundaries, but specifically health-related education is included. So are actions intended chiefly to improve health indirectly by influencing how non-health systems function – for example, actions to increase girls' school enrolment or change the curriculum to make students better future caregivers and consumers of health care.

### Box 1.1 Poverty, ill-health and cost-effectiveness

The series of global estimates of the burden of disease do not distinguish between rich and poor, but an approximate breakdown can be derived by ranking countries by per capita income, aggregating from the lowest and highest incomes to form groups each constituting 20% of the world's population, and studying the distribution of deaths in each group, by age,<sup>1</sup> cause and sex.<sup>2</sup> These estimates show that in 1990, 70% of all deaths and fully 92% of deaths from communicable diseases in the poorest quintile were "excess" compared to the mortality that would have occurred at the death rates of the richest quintile. The figures for total losses of disability-adjusted life

years (DALYs) were similar, with a larger contribution from noncommunicable diseases. The large difference between the effects of communicable and noncommunicable diseases reflects the concentration of deaths and DALYs lost to communicable diseases among the global poor: about 60% of all ill-health for the poor versus 8–11 % among the richest quintile. This is strongly associated with differences in the age distribution of deaths: just over half of all deaths among the poor occur before 15 years of age, compared to only 4% among the rich. The difference between the poor and the rich is large even in a typical high-mortality African country, and much greater in a typical lower-mortality Latin American

country, where deaths at early ages have almost been eliminated among the wealthy.

There are relatively cost-effective interventions available against the diseases that account for most of these rich–poor differences, and particularly to combat deaths and health losses among young children.<sup>3</sup> Interventions costing an estimated \$100 or less per DALY saved could deal with 8 or 9 of the 10 leading causes of ill-health under the age of 5 years, and 6 to 8 of the 10 main causes between the ages of 5 and 14 years. All of these are either communicable diseases or forms of malnutrition. Death and disability from these causes is projected to decline rapidly by 2020, roughly equalizing the health damage from

communicable and noncommunicable diseases among the poor. If the projected rate of decline of communicable disease damage could be doubled, the global rich would gain only 0.4 years of life expectancy, but the global poor would gain an additional 4.1 years, narrowing the difference between the two groups from 18.4 to 13.7 years. Doubling the pace of reduction of noncommunicable disease damage, in contrast, would preferentially benefit the well-off as well as costing considerably more. The association between poverty and cost-effectiveness is only partial, and probably transitory, but in today's epidemiological and economic conditions it is quite strong.

<sup>1</sup> Gwatkin DR. *The current state of knowledge about how well government health services reach the poor: implications for sector-wide approaches.* Washington, DC, The World Bank, 5 February 1998 (discussion draft).

<sup>2</sup> Gwatkin DR, Guillot M. *The burden of disease among the world's poor: current situation, future trends, and implications for policy.* Washington, DC, Human Development Network of The World Bank, 2000.

<sup>3</sup> *World development report 1993 – Investing in health.* New York, Oxford University Press for The World Bank, 1993: Tables B.6 and B.7.

This way of defining a system does not imply any particular degree of integration, nor that anyone is in overall charge of the activities that compose it. In this sense, every country has a health system, however fragmented it may be among different organizations or however unsystematically it may seem to operate. Integration and oversight do not determine the system, but they may greatly influence how well it performs.

Unfortunately, nearly all the information available about health systems refers only to the provision of, and investment in, health services: that is, the health *care* system, including preventive, curative and palliative interventions, whether directed to individuals or to populations. In most countries, these services account for the great bulk of employment, expenditure and activity that would be included in a broader notion of the health system, so it might seem that little is lost in concentrating on a narrower definition that fits the existing data. Those data have required great efforts to collect – and this report further offers several kinds of information and analysis, such as estimates of life expectancy adjusted for time lived with disability, assessments of how well health systems treat patients, national health accounts, and estimates of household contribution to financing.

Nonetheless, efforts are needed to quantify and assess those activities implied by the wider definition, so as to begin to gauge their relative cost and effectiveness in contributing to the goals of the system. To take one example, in the United States between 1966 and 1979 the introduction of a variety of safety features in automobile design (laminated windshields, collapsible steering columns, interior padding, lap and shoulder belts, side marker lights, head restraints, leak resistant fuel systems, stronger bumpers, increased side door strength and better brakes) helped reduce the vehicle accident fatality rate per mile travelled by 40%. Only three of these innovations added more than \$10 to the price of a car, and in total they accounted for only 2% of the average price increase during 1975–1979 (3). From 1975 to 1998, seat belts saved an estimated 112 000 lives in the United States, and total traffic fatalities continued to fall. The potential health gains were even greater: in 1998 alone, 9000 people died because they did not use their belts (4).

The potential savings in other countries are very large. Road traffic accidents are increasing rapidly in poor countries and are projected to move from the ninth to third place in the worldwide ranking of burden of ill-health by the year 2020. Even in many middle income countries, the fatality rates per head or per vehicle mile are much higher than in the United States (5). Sub-Saharan Africa has the world's highest rate of fatalities per vehicle. The cost of improving vehicles may be high, relative to expenditure on health care, in low and middle income countries, so the effect of including such activities in the definition of the health system may be greater. Unsafe roads also contribute greatly to the vehicular toll in poorer countries, and the cost of improving roads could be much larger than the cost of making cars safer. But behavioural changes such as using seat belts once installed, and respecting speed limits, are nearly costless and could save many lives; they are very likely to be more cost-effective than treatment of crash victims.

Where information corresponding to a broader definition of health systems is not available, this report necessarily uses the available data that match the notion of the health care system. Even by this more limited definition, health systems today represent one of the largest sectors in the world economy. Global spending on health care was about \$2985 billion (thousand million) in 1997, or almost 8% of world gross domestic product (GDP), and the International Labour Organisation estimates that there were about 35 million health workers worldwide a decade ago, while employment in health services now is likely to be substantially higher. These figures reflect how what was for thousands of years a basic, private relationship – in which one person with an illness was looked after by family mem-

bers or religious caregivers, or sometimes paid a professional healer to treat him or her – has expanded over the past two centuries into the complex network of activities that now comprise a health system.

More than simple growth, the creation of modern health systems has involved increasing differentiation and specialization of skills and activities. It has also involved an immense shift in the economic burden of ill-health. Until recently, most of that burden took the form of lost productivity, as people died young or became and remained too sick to work at full strength. The cost of health care accounted for only a small part of the economic loss, because such care was relatively cheap and largely ineffective. Productivity losses are still substantial, especially in the poorest countries, but success in prolonging life and reducing disability has meant that more and more of the burden is borne by health systems. This includes the cost of drugs – for controlling diabetes, hypertension, and heart disease, for example – that allow people to stay active and productive. Part of the growth in resources used by health systems is a transfer from other ways of paying for the economic damage due to illness and early death.

The resources devoted to health systems are very unequally distributed, and not at all in proportion to the distribution of health problems. Low and middle income countries account for only 18% of world income and 11% of global health spending (\$250 billion or 4% of GDP in those countries). Yet 84% of the world's population live in these countries, and they bear 93% of the world's disease burden. These countries face many difficult challenges in meeting the health needs of their populations, mobilizing sufficient financing in an equitable and affordable manner, and securing value for scarce resources.

Today in most developed countries – and many middle income countries – governments have become central to social policy and health care. Their involvement is justified on the grounds of both equity and efficiency. However, in low income countries – where total public revenues for all uses are scarce (often less than 20% of GDP) and institutional capacity in the public sector is weak – the financing and delivery of health services is largely in the hands of the private sector. In many of these countries, large segments of the poor still have no access to basic and effective care.

## WHAT DO HEALTH SYSTEMS DO?

For rich and poor alike, health needs today are very different from those of 100 or even 50 years ago. There are growing expectations of access to health care in some form, and growing demands for measures to protect the sick, and their families, against the financial costs of ill-health. The circle in which health systems are required to function has been pushed yet wider by raised awareness of the impact on health of developments such as industrialization, road transport, environmental damage and the globalization of trade. People also now turn to health systems for help with a much wider variety of problems than before – not just for the relief of pain and treatment of physical limitations and emotional disorders but for advice on diet, child-rearing and sexual behaviour that they used to seek from other sources.

People typically come into direct contact with a health system as patients, attended by providers, only once or twice a year. More often their contact is as consumers of non-prescription medications and as recipients of health-related information and advice. They meet the system as contributors to paying for it, knowingly every time they buy care out of pocket or pay insurance premiums or social security contributions, and unknowingly whenever they pay taxes that are used in part to finance health. It matters very much how the

system treats people's health needs and how it raises revenues from them, including how much protection it offers them from financial risk. But it also matters how it responds to their expectations. In particular, people have a right to expect that the health system will treat them with individual dignity. So far as possible, their needs should be promptly attended to, without long delays in waiting for diagnosis and treatment – not only for better health outcomes but also to respect the value of people's time and to reduce their anxiety. Patients also often expect confidentiality, and to be involved in choices about their own health, including where and from whom they receive care. They should not always be expected passively to receive services determined by the provider alone.

In summary, health systems have a responsibility not just to improve people's health but to protect them against the financial cost of illness – and to treat them with dignity. As is discussed in more detail in Chapter 2, health systems thus have three fundamental objectives. These are:

- improving the health of the population they serve;
- responding to people's expectations;
- providing financial protection against the costs of ill-health.

Because these objectives are not always met, public dissatisfaction with the way health services are run or financed is widespread, with accounts of errors, delays, rudeness, hostility and indifference on the part of health workers, and denial of care or exposure to calamitous financial risks by insurers and governments, on a grand scale.

Because better health is the most important objective of a health system, and because health status is worse in poor populations, one might assume that for a low income country, improving health is all that matters. Concern for the non-health outcomes of the system, for fairly sharing the burden of paying for health so that no one is exposed to great financial risk, and attending to people's wishes and expectations about how they are to be treated, would then be considered luxuries, gaining in importance only as income rises and health improves. But this view is mistaken, for several reasons. Poor people, as indicated earlier, need financial protection as much as or more than the well-off, since even small absolute risks may have catastrophic consequences for them. And the poor are just as entitled to respectful treatment as the rich, even if less can be done for them materially. Moreover, pursuing the objectives of responsiveness and financial protection does not necessarily take substantial resources away from activities to improve health. Much improvement in how a health system performs with respect to these responsibilities may often be had at little or no cost. So all three objectives matter in every country, independently of how rich or poor it is or how its health system is organized. Better ways of achieving these objectives, treated in later chapters, are similarly relevant for all countries and health systems, although the specific implications for policy will vary according to income level and the cultural and organizational features of the system.

## WHY HEALTH SYSTEMS MATTER

The contribution that health systems make to improving health has been examined much more closely than how well they satisfy the other two objectives mentioned above, for which there is little comparable information and analysis. This report therefore develops measures corresponding to all three objectives, for assessing how systems perform. Even the contribution that health systems make to improved health is difficult to judge, because different kinds of evidence seem to give conflicting answers. At the level of interventions

against particular diseases or conditions, there is now substantial and growing evidence that large improvements in health can be achieved at reasonable cost, for individuals and for large populations (6). Such data are the basis for estimates that in poor countries, roughly one-third of the disease burden in 1990 might be averted at a total cost per person of only \$12 (7).

Even without progress in fundamental science, changes in the way currently available interventions are organized and delivered can reverse the spread of an epidemic and dramatically reduce the cost of saving a life. For example, in the Brazilian Amazon, greater emphasis on early malaria case detection and treatment, together with more focused efforts on mosquito control, turned around an epidemic and cut the cost of saving a life by case prevention from nearly \$13 000 to only about \$2000 (8).

At the level of overall progress in health, as reported in *The world health report 1999*, the generation and utilization of knowledge – that is, scientific and technical progress – explained almost half of the reduction in mortality between 1960 and 1990 in a sample of 115 low and middle income countries, while income growth explained less than 20% and increases in the educational level of adult females less than 40%. Such estimates summarize progress in developing and applying interventions of many kinds against a large number of diseases. Prominent among these are antimalaria and immunization programmes, and the increasing use of antibiotics for the treatment of respiratory and other infectious diseases. Since it is the health system that develops and applies those interventions, two kinds of evidence, one detailed and the other aggregated, indicate clearly that health systems not only can but do make a large difference to health.

Taking a narrower focus on diseases for which there are effective treatments, numerous studies beginning in the 1970s (9, 10) have consistently found that preventable deaths, that is “deaths due to causes amenable to medical care” have fallen at a faster rate than other deaths. Similarly, a comparison of death rate differences between western Europe and formerly communist countries of eastern Europe attributed 24% of the difference in male life expectancy and 39% of that in female life expectancy to the availability of modern medical care. Such care is not guaranteed simply by the existence of medical facilities (11).

At the same time, other evidence seems to show that health systems make little or no difference. This emerges from some other comparisons across countries rather than through time. Often these show that while per capita income is strongly related to some measure of health status – as are other factors such as female education, income inequality or cultural characteristics – there is little independent connection with inputs such as doctors or hospital beds (12), with total health expenditure (13), with expenditure only on conditions amenable to medical care (14), or with public spending on health (15). It is not surprising to find that these relations are weak in rich countries, since many causes of death and disability are already controlled and there are many different ways to spend health system resources, with quite varying effects on health status. But health system expenditure often seems to make little difference even in poor countries with high infant and child mortality, which it should be a priority to reduce.

Furthermore, health systems make costly, even fatal mistakes far too frequently. In the United States alone, medical errors in hospitals cause at least 44 000 needless deaths a year, with another 7000 occurring as a result of mistakes in prescribing or using medication, making these errors more deadly than such killers as motor vehicle accidents, breast cancer and AIDS (16). The economic cost of these mistakes is at least \$17 billion, of which health care costs are more than half. And even when no one makes errors, patients often acquire

new infections in hospital, and the massive use of antibiotics promotes pathogen resistance to them, so that some part of ill-health is caused by the very efforts to treat it.

These conflicting kinds of evidence can be reconciled in two ways: first, by noting that while health systems account for much health progress through time, that progress is far from uniform among countries at any one time, even among countries with similar levels of income and health expenditure; and second, by recognizing that the errors of the system diminish but do not offset the good it accomplishes. Nonetheless, “there is an enormous gap between the apparent potential of public spending to improve health status and the actual performance” (15), and the same is doubtless true of resource use in general. One measure of that gap is that many deaths of children under 5 years of age could be averted for \$10 or less, as estimated from cost-effectiveness studies of particularly valuable interventions, but the average actual expenditure in poor countries per death prevented, as estimated from the overall relation between spending and mortality, is \$50 000 or more. The overall relation between child mortality and income implies that in a poor country of two million population, total income would have to rise by roughly \$1 million in order to avert a single death. This is several orders of magnitude higher than the average health expenditure needed to save a life. Per capita, these numbers imply health expenditure of only \$0.025 versus an income increase of \$0.50. Income differences may explain more of health variation among countries than do differences in health expenditure. But raising income is not on that account a cheaper or easier way to improve health.

Concerning the more distant past, historians debate whether declines in mortality rates in some European and Latin American countries in the 19th and 20th centuries owe more to such factors as an improving diet and other socioeconomic progress than to personal medical care. But health systems, defined broadly, include all of the non-personal, population-based or public health interventions such as the promotion of healthy lifestyles, insecticide spraying against vector-borne diseases, anti-tobacco campaigns and the protection of food and water. So even if personal services accounted for very little health gain until recently, the health system as defined in this report began to make a large difference more

### Box 1.2 Health knowledge, not income, explains historical change in urban–rural health differences

In the first half of the 19th century, life expectancy was much shorter in London and Paris, respectively, than in the rural areas of England and Wales or of France; a similar difference prevailed between the urban and rural areas of Sweden in the first decades of the 20th century. Large cities were unhealthy because unclean personal habits did more to spread disease when people were crowded together and because garbage and even excrement were accumulated, drawing flies and rodents and contaminating the air and water.<sup>1</sup> Pollution was wors-

ened by burning soft coal and by discharges from factories.

Crowding and poverty produce many of the same problems in the large cities of poor countries today, which typically have more polluted air and water than urban areas in richer countries. Vehicular exhaust, unknown a century ago, is already a major health threat in such areas as Delhi and Mexico City. Rapid growth has made it hard to expand such services as piped water, sewerage facilities and garbage collection fast enough to keep pace. In slum areas, even if safe water is available, many households have no access to sani-

tary waste disposal, and much garbage is simply dumped or burned in the open. Nonetheless the health consequences are not so severe as in European cities 150 years ago. On one hand, increased knowledge of how diseases are caused and transmitted has led to valiant efforts to reduce contamination, control disease vectors and educate the population to take better care of their health. On the other hand, even very poor urban dwellers now have better access to effective personal health care than much of the rural population, adding to the inducements to migrate to the city. Slum

residents in Lima, for example, are as likely to immunize their children and to take them for medical care when sick as residents of better-off neighbourhoods, and much more likely to do so than people living in Peru's mountainous interior.<sup>2</sup> Both the public health and the personal care interventions have contributed to reversing the urban–rural differences in health status; better health among urban populations is due more to the application of improved knowledge than to higher incomes in cities.

<sup>1</sup> Easterlin RA. *How beneficent is the market? A look at the modern history of mortality*. Los Angeles, University of Southern California, 1998 (unpublished paper).

<sup>2</sup> Musgrove P. Measurement of equity in health. *World Health Statistics Quarterly*, 1986, 39(4).

than a century ago, chiefly through improvements in urban sanitation and personal hygiene. These changes – removing excrement and garbage, protecting water supplies, and washing one’s hands – happened because of more understanding of how diseases are *spread*, even before there was any useful knowledge of how they are *caused*. Some improved individual hygienic practices are centuries old, while collective measures are generally more recent. Growth in income alone would not have improved health under the conditions of the time, and may even have worsened it because of urban filth and crowding; similar conditions often prevail in the cities of poor countries today, but the threat to health is better controlled (see Box 1.2).

So health systems *are* valuable and important, but they could accomplish much more with the available understanding of how to improve health. The failings which limit performance do not result primarily from lack of knowledge but from not fully applying what is already known: that is, from systemic rather than technical failures. This is true even of most medical errors, because “the problem is not bad people; the problem is that the system needs to be made safer” (16). How to measure current performance and how to achieve the potential improvements in it are the subject of this report. Research to expand knowledge is crucial in the long run, as progress over the last two centuries shows; in the short run, much could be accomplished by wider and better application of existing knowledge. This can improve health more quickly than continued and more equally distributed socio-economic progress, important as that is. The next sections discuss how modern health systems arose, and how they have been repeatedly subjected to reforms intended to make them work better in one way or another.

## HOW MODERN HEALTH SYSTEMS EVOLVED

Health systems of some sort have existed for as long as people have tried deliberately to protect their health and treat diseases. Throughout the world, traditional practices based on herbal cures, often integrated with spiritual counselling, and providing both preventive and curative care, have existed for thousands of years, and often coexist today with modern medicine. Many of them are still the treatment of choice for some health conditions, or are resorted to because modern alternatives are not understood or trusted, or fail, or are too expensive. Traditional Chinese medicine can be traced back more than 3000 years, and still plays a huge role in the Chinese health system, as do its equally ancient equivalents in the Indian sub-continent and similar systems of belief and practice among indigenous African and American peoples. But until the modern growth of knowledge about disease, there were few cures for ailments and little effective prevention of disease.

With rare exceptions, even in industrialized countries, organized health systems in the modern sense, intended to benefit the population at large, barely existed a century ago. Although hospitals have a much longer history than complete systems in many countries, few people living 100 years ago would ever visit one – and that remains true for many millions of the poor today. Until well into the 19th century they were for the most part run by charitable organizations, and often were little more than refuges for the orphaned, the crippled, the destitute or the insane. And there was nothing like the modern practice of referrals from one level of the system to another, and little protection from financial risk apart from that offered by charity or by small-scale pooling of contributions among workers in the same occupation.

Towards the close of the 19th century, the industrial revolution was transforming the lives of people worldwide. At the same time societies began to recognize the huge toll of

death, illness and disability occurring among workforces, whether from infectious diseases which killed many thousands during the construction of the Panama Canal or from industrial accidents and exposures. Once it was realized that mosquitoes transmitted malaria and yellow fever, control of the insects' breeding-sites became part of prevention efforts that also translated into benefits for surrounding communities. In addition to the human costs, the toll of illness and death meant great losses in productivity. In response, company owners began providing medical services to treat their workers. As the importance of clean water and sanitation became better understood, they also improved workers' basic living conditions. Wars were another influence – the American Civil War showed that soldiers on both sides were more likely to be killed by disease than by the enemy. The same message came home from the Crimean and Boer wars.

About the same time, workers' health was becoming a political issue in some European countries, but for quite different reasons. Bismarck, Chancellor of Germany, reasoned that government take-over of labour unions' sickness funds would remove a source of their support at a moment when socialist workers' movements were gaining strength, and also increase workers' economic security (17). Thus, in 1883, Germany enacted a law requiring employer contributions to health coverage for low-wage workers in certain occupations, adding other classes of workers in subsequent years. This was the first example of a state-mandated social insurance model. The popularity of this law among workers led to the adoption of similar legislation in Belgium in 1894 and Norway in 1909. Until Britain followed suit in 1911, medical care for British wage-earners tended to be paid for by their subscriptions to trade unions or friendly societies, which in turn paid the providers. But only the worker, and not his family, had such coverage.

In the late 1800s, Russia had begun setting up a huge network of provincial medical stations and hospitals where treatment was free and supported by tax funds. After the Bolshevik revolution in 1917, it was decreed that free medical care should be provided for the entire population, and the resulting system was largely maintained for almost eight decades. This was the earliest example of a completely centralized and state-controlled model.

The influence of the German model began to spread outside Europe after the First World War. In 1922, Japan added health benefits to the other benefits for which workers were eligible, building on its tradition of managerial paternalism. In 1924, Chile brought all covered workers under the umbrella of a Ministry of Labour scheme. By 1935, a total of 90% of Denmark's population was covered by work-related health insurance. Social insurance was introduced in the Netherlands during the country's occupation in the Second World War.

Not least among its effects, the Second World War damaged or virtually destroyed health infrastructures in many countries and delayed their health system plans. Paradoxically, it also paved the way for the introduction of some others. Wartime Britain's national emergency service to deal with casualties was helpful in the construction of what became, in 1948, the National Health Service, perhaps the most widely influential model of a health system. The Beveridge Report of 1942 (18) had identified health care as one of the three basic prerequisites for a viable social security system. The government's White Paper of 1944 stated the policy that "Everybody, irrespective of means, age, sex or occupation shall have equal opportunity to benefit from the best and most up-to-date medical and allied services available", adding that those services should be comprehensive and free of charge and should promote good health, as well as treating sickness and disease. New Zealand had already become, in 1938, the first country to introduce a national health service. Almost simultaneously, Costa Rica laid the foundation for universal health insurance in 1941. In

Mexico, the Institute of Social Security and the Ministry of Health were both founded in 1943. A scheme for a national health service broadly similar to the British model was proposed in South Africa in 1944, comprising free health care and a network of community centres and general practitioners as part of a referral system, but was not implemented (19).

In the immediate post-war period, Japan and the Soviet Union also extended their limited national systems to cover most or all of the population, as did Norway and Sweden, Hungary and other communist states in Europe, and Chile. As former colonies gained independence, they also tried to adopt modern, comprehensive systems with heavy state participation. India developed ambitious five-year development plans for a health system, based on the Bhore Report of 1946 (20). The factors which made this period of system-building and expansion possible included realization of the power of the modern state, post-war movements towards reconciliation, stability and reconstruction, and collective solidarity stemming from the war effort. Newly acquired citizenship and the belief in a relatively effective and benevolent state which could promote development of all kinds led to a social and political environment in which “classical universalism”, the concept of free access to all kinds of health care for all, could take root.

Today’s health systems are modelled to varying degrees on one or more of a few basic designs that emerged and have been refined since the late 19th century. One of these aims to cover all or most citizens through mandated employer and employee payments to insurance or sickness funds, while providing care through both public and private providers. The earliest such social insurance systems usually evolved from small, initially voluntary, associations; later versions have sometimes been created *ex nihilo* by public action. Another, slightly more recent, model centralizes planning and financing, relying primarily on tax revenues and on public provision. Resources are traditionally distributed by budgets, sometimes on the basis of fixed ratios between populations and health workers or facilities. In a third model, state involvement is more limited but still substantial, sometimes providing coverage only for certain population groups and giving way for the rest of the populace to largely private finance, provision and ownership of facilities. Relatively pure examples, in which one or another model accounts for the bulk of resources or provision, are found mostly in rich countries; health systems in middle income countries, notably in Latin America, tend to be a mixture of two or even all three types (21). Much debate has centred on whether one way of organizing a health system is better than another, but what matters about a system’s overall structure is how well it facilitates the performance of its key functions.

## THREE GENERATIONS OF HEALTH SYSTEM REFORM

During the 20th century, there have been three overlapping generations of health system reforms. They have been prompted not only by perceived failures in health but also by a quest for greater efficiency, fairness and responsiveness to the expectations of the people that systems serve. The first generation saw the founding of national health care systems, and the extension to middle income nations of social insurance systems, mostly in the 1940s and 1950s in richer countries and somewhat later in poorer countries. By the late 1960s, many of the systems founded a decade or two earlier were under great stress. Costs were rising, especially as the volume and intensity of hospital-based care increased in developed and developing countries alike. Among systems that were nominally universal in coverage, health services still were used more heavily by the better-off, and efforts to reach the poor were often incomplete. Too many people continued to depend on their own resources to pay for health, and could often get only ineffective or poor quality care.

These problems were apparent, and increasingly acute, in poorer countries. Colonial powers in Africa and Asia, and governments in Latin America, had established health services that for the most part excluded indigenous populations. For example, where a European model of health care was implemented in the countries of Africa under British administration, it was primarily intended for colonial administrators and expatriates, with separate or second class provision made – if at all – for Africans. Charitable missions and public health programmes were relied on to provide some care for the majority, much as in parts of Europe. In these former colonies and low income countries, the health system had therefore never been able to deliver even the most basic services to people in rural areas. Health facilities and clinics had been built, but primarily in urban areas. In most developing countries, major urban hospitals received around two-thirds of all government health budgets, despite serving just 10% to 20% of the population. Studies of what hospitals actually did revealed that half or more of all inpatient spending went towards treating conditions that could often have been managed by ambulatory care, such as diarrhoea, malaria, tuberculosis and acute respiratory infections (22).

There was, therefore, a need for radical change that would make systems more cost-efficient, equitable, and accessible. A second generation of reforms thus saw the promotion of primary health care as a route to achieving affordable universal coverage. This approach reflected experience with disease control projects in the 1940s in countries such as South Africa, the Islamic Republic of Iran, and former Yugoslavia. It also built on the successes and experiments of China, Cuba, Guatemala, Indonesia, Niger, the United Republic of Tanzania, and Maharashtra State in India (23). Some of these countries, and others such as Costa Rica and Sri Lanka, achieved very good health outcomes at relatively little cost, adding 15 to 20 years to life expectancy at birth in a span of just two decades. In each case, there was a very strong commitment to assuring a minimum level for all of health services, food and education, along with an adequate supply of safe water and basic sanitation. These were the key elements, along with an emphasis on public health measures relative to clinical care, prevention relative to cure, essential drugs, and education of the public by community health workers. By adopting primary health care as the strategy for achieving the goal of “Health for All” at the Joint WHO/UNICEF International Conference on Primary Health Care held at Alma-Ata, USSR (now Almaty, Kazakhstan) in 1978, WHO reinvigorated efforts to bring basic health care to people everywhere.

The term “primary” quickly acquired a variety of connotations, some of them technical (referring to the first contact with the health system, or the first level of care, or simple treatments that could be delivered by relatively untrained providers, or interventions acting on primary causes of disease) and some political (depending on multisectoral action or community involvement). The multiplicity of meanings and their often contradictory implications for policy help explain why there is no one model of primary care, and why it has been difficult to follow the successful examples of the countries or states that provided the first evidence that a substantial improvement in health could be achieved at affordable cost. There was a substantial effort in many countries to train and use community health workers who could deliver basic, cost-effective services in simple rural facilities to populations that previously had little or no access to modern care. In India, for example, such workers were trained and placed in over 100 000 health posts, intended to serve nearly two-thirds of the population.

Despite these efforts, many such programmes were eventually considered at least partial failures. Funding was inadequate; the workers had little time to spend on prevention and community outreach; their training and equipment were insufficient for the problems

they confronted; and quality of care was often so poor as to be characterized as “primitive” rather than “primary”, particularly when primary care was limited to the poor and to only the simplest services. Referral systems, which are unique to health services and necessary to their proper performance, have proved particularly difficult to operate adequately (24). Lower level services were often poorly utilized, and patients who could do so commonly bypassed the lower levels of the system to go directly to hospitals. Partly in consequence, countries continued to invest in tertiary, urban-based centres.

In developed countries, primary care has been better integrated into the whole system, perhaps because it has been more associated with general and family medical practice, and with lower-level providers such as nurse practitioners and physician assistants. Greater reliance on such practitioners forms the core of many developed countries’ current reform agendas. Managed care, for example, revolves to a large extent around the strengthening of primary care and the avoidance of unnecessary treatment, especially hospitalization.

The approach emphasized in the primary health care movement can be criticized for giving too little attention to people’s *demand* for health care, which is greatly influenced by perceived quality and responsiveness, and instead concentrating almost exclusively on their presumed *needs*. Systems fail when these two concepts do not match, because then the supply of services offered cannot possibly align with both. The inadequate attention to demand is reflected in the complete omission of private finance and provision of care from the Alma-Ata Declaration, except insofar as community participation is construed to include small-scale private financing.

Poverty is one reason why needs may not be expressed in demand, and that can be resolved by offering care at low enough cost, not only in money but also in time and non-medical expenses. But there are many other reasons for mismatches between what people need and what they want, and simply providing medical facilities and offering services may do nothing to resolve them. In general, both the first-generation and second-generation reforms have been quite supply-oriented. Concern with demand is more characteristic of changes in the third generation currently under way in many countries, which include such reforms as trying to make “money follow the patient” and shifting away from simply giving providers budgets, which in turn are often determined by supposed needs.

If the organizational basis and the quality of primary health care often failed to live up to their potential, much of the technical footing remains sound and has undergone continuous refinement. This development can be sketched as a gradual convergence towards what WHO calls the “new universalism”– high quality delivery of essential care, defined mostly by the criterion of cost-effectiveness, for everyone, rather than all possible care for the whole population or only the simplest and most basic care for the poor (see Figure 1.1).

**Figure 1.1 Coverage of population and of interventions under different notions of primary health care**

	Population covered	
Interventions included	Only the poor	Everyone
"Basic" or simple	"Primitive" health care	Original concept
"Essential" and cost-effective	"Selective" primary health care	New universalism
Everything medically useful	(Never seriously contemplated)	Classical universalism

Adapted from Frenk J. *Building on the legacy: primary health care and the new policy directions at WHO*. Address to the American Public Health Association, Chicago, IL, 8 November 1999.

The notions that health and nutrition interventions can make a substantial difference to the health of large populations (25) and of obtaining “good health at low cost” (26) by selectively concentrating efforts against diseases that account for large, avoidable burdens of ill-health, are the basis for packages of interventions, variously called “basic” or “essential” or “priority”, that have been developed in several countries from epidemiological information and estimates of cost-effectiveness of interventions (27, 28). And the common failures in diagnosis and treatment due to inadequate training and excessive separation among disease control efforts have led to the development of clusters of interventions and more thorough training to support their delivery, most notably in the integrated management of childhood illness (29).

This evolution also implies an emphasis on public or publicly guaranteed and regulated finance, but not necessarily on public delivery of services. And it implies explicit choice of priorities among interventions, respecting the ethical principle that it may be necessary and efficient to ration services but that it is inadmissible to exclude whole groups of the population. However, it is easier to define a set of interventions that would preferentially benefit the poor if fully applied to the population, than it is to assure either that most of the poor actually do benefit, or that most of the beneficiaries are poor. Government health care services, although usually intended to reach the poor, often are used more by the rich. In 11 countries for which the distribution of benefits has been calculated from the distribution of public expenditure and utilization rates, the poorest quintile of the population never accounts for even its equal share (20%), and in seven of those countries the richest quintile takes 29% to 33% of the total benefit. This pro-rich bias is due largely to disproportionate use of hospital services by the well-off, who (with one exception) always account for at least 26% of the overall benefit. The distribution of primary care is almost always more beneficial to the poor than hospital care is, justifying the emphasis on the former as the way to reach the worst-off. Even so, the poor sometimes obtain less of the benefit of primary care than the rich (30). The poor often obtain much of their personal ambulatory care – which accounts for the bulk of their use of the health system and their out-of-pocket expenditure, and offers the greatest opportunity for further health gains – from private providers (31), and those services may be either more or less pro-poor than the care offered by the public sector.

The ideas of responding more to demand, trying harder to assure access for the poor, and emphasizing financing, including subsidies, rather than just provision within the public sector, are embodied in many of the current third-generation reforms. These efforts are more difficult to characterize than earlier reforms, because they arise for a greater variety of reasons and include more experimentation in approach. In part, they reflect the profound political and economic changes that have been taking place in the world. By the late 1980s, the transformation from communist to market-oriented economies was under way in China, central Europe, and the former Soviet Union. Heavy-handed state intervention in the economy was becoming discredited everywhere, leading to widespread divestiture of state enterprises, promotion of more competition both internally and externally, reduction in government regulation and control, and in general, much more reliance on market mechanisms. Ideologically, this meant greater emphasis on individual choice and responsibility. Politically, it meant limiting promises and expectations about what governments should do, particularly via general revenues, to conform better to their actual financial and organizational capacities.

Health systems have not been immune from these large-scale changes. One consequence has been a greatly increased interest in explicit insurance mechanisms, including

privately financed insurance. Reforms including such changes have occurred in several Asian countries, universal health insurance being introduced to different degrees in the Republic of Korea, Malaysia, Singapore and China (Province of Taiwan). Reforms to consolidate, extend or merge insurance coverage for greater risk-sharing have also occurred in Argentina, Chile, Colombia and Mexico, and a mixture of insurance and out-of-pocket health care has replaced much of the public system throughout the former communist countries. In developed countries which already had essentially universal coverage, usually less drastic changes have taken place in how health care is financed. But there have been substantial changes in who determines how resources are used, and in the arrangements by which funds are pooled and paid to providers. General practitioners and primary care physicians, as ‘gatekeepers’ to the health system, have sometimes been made accountable “not only for their patients’ health but also for the wider resource implications of any treatments prescribed. In some countries this role has been formalised through establishing ‘budget holding’ for general practitioners and primary care physicians, for example, through general practice ‘fund holding’ in the UK, Health Maintenance Organizations in the USA, and Independent Practice Associations in New Zealand” (32). And in the United States, there has been a great shift of power from providers to insurers, who now largely control the access of doctors and patients to one another.

## FOCUSING ON PERFORMANCE

This report does not analyse the variety of current reform efforts and proposals in detail, nor offer a model of how to construct or reconstruct a health system. The world is currently experimenting with many variants, and there is no clearly best way to proceed. But there do seem to be some clear conclusions about the organizations, rules and incentives that best help a health system to use its resources to achieve its goals; these are the subject of Chapter 3. How much can be accomplished with currently available resources – people, buildings, equipment and knowledge – depends greatly on the past investment and training that created those resources. And mistakes in investment have long-lasting consequences. The questions of how best to create resources, and what mistakes to avoid, are the subject of Chapter 4. There are comparable conclusions about what is desirable in the financing of the system; these are treated in Chapter 5. Finally, the health system as a whole needs comprehensive oversight, to stay directed to its goals and to ensure that the tasks of financing, investing and delivering services are adequately carried out. Suggestions concerning this more general function are developed in Chapter 6. These subjects are emphasized partly because so much reform today aims to change such aspects, rather than simply expanding supply or determining which interventions to offer. And all changes, to be justified, need to improve the performance of the system.

How then can the potential of health systems be fulfilled? How can they perform better, so that besides protecting health, they respond to people’s expectations, and protect them financially against the costs of ill-health? Chapter 2 sets out a framework for assessing health system performance and understanding the factors that contribute to it in the four key areas treated in subsequent chapters: providing services, developing the resources – human, material and conceptual – required for the system to work, mobilizing and channelling financing, and ensuring that the individuals and organizations that compose the system act as good stewards of the resources and trust given to their care.

## REFERENCES

1. Miller J. *The body in question*. New York, Random House, 1978: 14.
2. Kleinman A. Concepts and a model for the comparison of medical systems as cultural systems. *Social Science and Medicine*, 1978, **12**: 85–93.
3. Claybrook J. Remarks at the Seventh Annual North Carolina Highway Safety Conference, 1980. On Bureau of Transportation Statistics, Transportation Research Board web site at <http://www4.nas.edu/trb/crp.nsf/>
4. US Department of Transportation. *Traffic safety facts 1998*. Washington, DC, National Highway Traffic Safety Administration, 1998.
5. OECD. *International road traffic and accident database*. Paris, Organisation for Economic Co-operation and Development, 1999.
6. Jamison DT et al. *Disease control priorities in developing countries*. New York, Oxford University Press for The World Bank, 1993.
7. *World development report 1993 – Investing in health*. New York, Oxford University Press for The World Bank, 1993: Table 5.3.
8. Akhavan D et al. Cost-effective malaria control in Brazil. Cost-effectiveness of a malaria control program in the Amazon Basin of Brazil, 1988–1996. *Social Science and Medicine*, 1999, **49**(10): 1385–99: Table 5.
9. Rutstein DD et al. Measuring the quality of medical care – a clinical method. *New England Journal of Medicine*, 1976, **294**(11): 582–588.
10. Charlton JR, Velez R. Some international comparisons of mortality amenable to medical intervention. *British Medical Journal*, 1986, **292**: 295–301.
11. Velkova A, Wolleswinkel-van den Bosch JH, Mackenbach JP. The East–West life expectancy gap: differences in mortality from conditions amenable to medical intervention. *International Journal of Epidemiology*, 1997, **26**(1): 75–84.
12. Cochrane AL, St Leger AS, Moore F. Health service ‘input’ and mortality ‘output’ in developed countries. *Journal of Epidemiology and Community Health*, 1978, **32**(3): 200–205.
13. Musgrove P. *Public and private roles in health: theory and financing patterns*. Washington, DC, The World Bank, 1996 (World Bank Discussion Paper No. 339).
14. Mackenbach JP. Health care expenditure and mortality from amenable conditions in the European Community. *Health Policy*, 1991, **19**: 245–255.
15. Filmer D, Pritchett L. The impact of public spending on health: does money matter? *Social Science and Medicine*, 1999, **49**(10): 1309–1323.
16. Kohn L, Corrigan J, Donaldson M, eds. *To err is human: building a safer health system*. Washington, DC, Institute of Medicine, National Academy of Sciences, 1999.
17. Taylor ASP. *Bismarck – the man and the statesman*. London, Penguin, 1995: 204.
18. *Social insurance and allied services. Report by Sir William Beveridge*. London, HMSO, 1942.
19. Savage M, Shisana O. Health service provision in a future South Africa. In: Spence J, ed. *Change in South Africa*. London, The Royal Institute of International Affairs, 1994.
20. Government of India. *Health Survey and Development Committee Report. Vol. 1–4*. New Delhi, Ministry of Health, 1946.
21. Londoño JL, Frenk J. Structured pluralism: towards an innovative model for health system reform in Latin America. *Health Policy*, 1997, **41**(1) :1–36.
22. Barnum H, Kutzin J. *Public hospitals in developing countries: resource use, cost, financing*. Baltimore, MD, The Johns Hopkins University Press, 1993.
23. Newell KN. *Health by the people*. Geneva, World Health Organization, 1975.
24. Sanders D et al. Zimbabwe’s hospital referral system: does it work? *Health Policy and Planning*, 1998, **13**: 359–370.
25. Gwatkin DR, Wilcox JR, Wray JD. The policy implications of field experiments in primary health nutrition care. *Social Science and Medicine*, 1980, **14**(2): 121–128.
26. Halstead SB, Walsh JA, Warren KS, eds. *Good health at low cost*. New York, Rockefeller Foundation, 1985.

27. **Bobadilla JL 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.
28. **Bobadilla JL.** *Searching for essential health services in low- and middle-income countries*. Washington, DC, Inter-American Development Bank, 1998.
29. **Tulloch J.** Integrated approach to child health in developing countries. *The Lancet*, 1999, **354**(Suppl. II): 16–20.
30. **Gwatkin DR.** *The current state of knowledge about how well government health services reach the poor: implications for sector-wide approaches*. Washington, DC, The World Bank, 5 February 1998 (discussion draft).
31. **Berman P.** The organization of ambulatory care provision: a critical determinant of health system performance in developing countries. *Bulletin of the World Health Organization*, 2000, **78**(6) (in press).
32. **Wilton P, Smith RD.** Primary care reform: a three country comparison of ‘budget holding’. *Health Policy*, 1998, **44**(2): 149–166.

## CHAPTER TWO

# *How Well do Health Systems Perform ?*

*Better health is unquestionably the primary goal of a health system. But because health care can be catastrophically costly and the need for it unpredictable, mechanisms for sharing risk and providing financial protection are important. A second goal of health systems is therefore fairness in financial contribution. A third goal – responsiveness to people’s expectations in regard to non-health matters – reflects the importance of respecting people’s dignity, autonomy and the confidentiality of information. WHO has engaged in a major exercise to obtain and analyse data in order to assess how far health systems in WHO Member States are achieving these goals for which they should be accountable, and how efficiently they are using their resources in doing so. By focusing on a few universal functions that health systems undertake, this report provides an evidence base to assist policy-makers improve health system performance.*

## 2

## HOW WELL DO HEALTH SYSTEMS PERFORM?

### ATTAINMENT AND PERFORMANCE

Assessing how well a health system does its job requires dealing with two large questions. The first is how to measure the outcomes of interest – that is, to determine what is achieved with respect to the three objectives of good health, responsiveness and fair financial contribution (*attainment*). The second is how to compare those attainments with what the system *should* be able to accomplish – that is, the best that could be achieved with the same resources (*performance*). Although progress is feasible against many of society’s health problems, some of the causes lie completely outside even a broad notion of what health systems are. Health systems cannot be held responsible for influences such as the distribution of income and wealth, any more than for the impact of the climate. But avoidable deaths and illness from childbirth, measles, malaria or tobacco consumption can properly be laid at their door. A fair judgement of how much health damage it should be possible to avoid requires an estimate of the best that can be expected, and of the least that can be demanded, of a system. The same is true of progress towards the other two objectives, although much less is known about them (1).

### GOALS AND FUNCTIONS

Better health is of course the *raison d’être* of a health system, and unquestionably its primary or defining goal: if health systems did nothing to protect or improve health there would be no reason for them. Other systems in society may contribute greatly to the population’s health, but not as their primary goal. For example, the education system makes a large difference to health, but its defining goal is to educate. Influence also flows the other way: better health makes children better able to learn, but that is not the defining purpose of the health system. In contrast, the goal of fair financing is common to all societal systems. This is obvious when the system is paid for socially, but it holds even when everything is financed purely by individual purchases. It is only the notion of fairness that may vary. “Getting what you pay for” is generally accepted as fair in market transactions, but seems much less fair where health services are concerned. Similarly, in any system, people have expectations which society regards as legitimate, as to how they should be treated, both physically and psychologically. Responsiveness is therefore always a social goal. Taking the education system as an example, fair financing means the right balance of contributions from households which do and those which do not have children in school, and enough subsidy that poor children are not denied schooling for financial reasons. Responsiveness includes respect for parents’ wishes for their children, and avoiding abuse or humiliation of the students themselves.

The health system differs from other social systems such as education, and from the markets for most consumer goods and services, in two ways which make the goals of fair financing and responsiveness particularly significant. One is that health care can be catastrophically costly. Much of the need for care is unpredictable, so it is vital for people to be protected from having to choose between financial ruin and loss of health. Mechanisms for sharing risk and providing financial protection are more important even than in other cases where people buy insurance, as for physical assets like homes or vehicles, or against the financial risk to the family of a breadwinner dying young. The other peculiarity of health is that illness itself, and medical care as well, can threaten people's dignity and their ability to control what happens to them more than most other events to which they are exposed. Among other things, responsiveness means reducing the damage to one's dignity and autonomy, and the fear and shame that sickness often brings with it.

Systems are often charged to be affordable, equitable, accessible, sustainable, of good quality, and perhaps to have many other virtues as well. However, desiderata such as accessibility are really a means to an end; they are instrumental rather than final goals. The more accessible a system is, the more people should utilize it to improve their health. In contrast, the goals of health, fair financing, and responsiveness are each intrinsically valuable. Raising the achievement of any goal or combination of goals, without lowering the attainment of another, represents an improvement. So if the achievement of these goals can be measured, then instrumental goals such as accessibility become unnecessary as proxy measures of overall performance; they are relevant rather as explanations of good or bad outcomes.

It is certainly true that financing that is more fairly distributed may contribute to better health, by reducing the risk that people who need care do not get it because it would cost too much, or that paying for health care leaves them impoverished and exposed to more health problems. And a system that is more responsive to what people want and expect can also make for better health, because potential patients are more likely to utilize care if they anticipate being treated well. Both goals therefore are partly instrumental, in that they promote improvements in health status. But they would be valuable even if that did not happen. That is, paying equitably for the system is a good thing in itself. So is assuring that people are treated promptly, with respect for their dignity and their wishes, and that patients receive adequate physical and affective support while undergoing treatment. The three goals are separable, as is often shown by people's unhappiness with a system even when the health outcomes are satisfactory.

Comparing how health systems perform means looking at what they achieve and at what they *do* – how they carry out certain *functions* – in order to achieve anything (2). These functions could be classified and related to system objectives in many different ways. For example, the “Public health in the Americas” initiative led by the Pan American Health Organization describes 12 different “essential functions”, and proposes between three and six sub-functions for each one (3). Many of these functions correspond to the task of stewardship which this report emphasizes, others to service provision and to resource generation. The four functions described in this chapter embrace these and other more specific activities. Figure 2.1 indicates how these functions – delivering personal and non-personal health services; raising, pooling and allocating the revenues to purchase those services; investing in people, buildings and equipment; and acting as the overall stewards of the resources, powers and expectations entrusted to them – are related to one another and to the objectives of the system. Stewardship occupies a special place because it involves oversight of all the other functions, and has direct or indirect effects on all the outcomes. Comparing the way these functions are actually carried out provides a basis for understanding

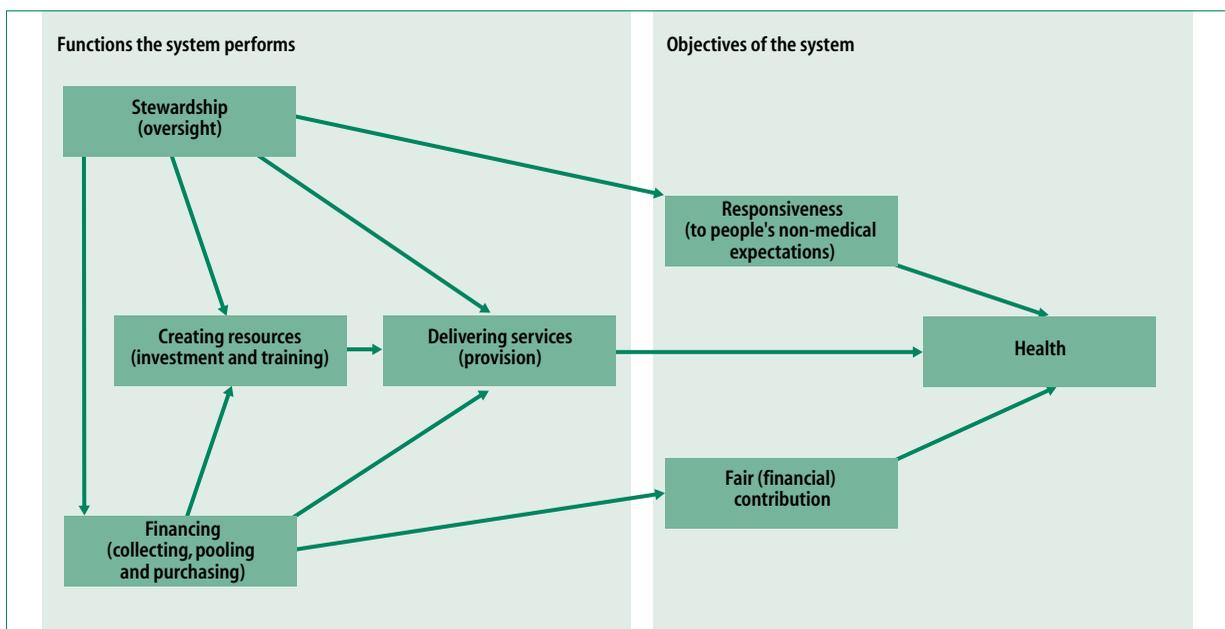
performance variations over time and among countries. Some evidence about these functions, and how they influence the attainment of the fundamental objectives in different health systems, is examined in the next four chapters.

In the view of most people, the health system is simply those providers and organizations which deliver personal medical services. Defining the health system more broadly means that the people and organizations which deliver medical care are not the whole system; rather, they exercise one of the principal functions of the system. They also share, sometimes appropriately and sometimes less so, in the other functions of financing, investment and stewardship. The question of who should undertake which functions is one of the crucial issues treated in later chapters.

It is common to describe the struggle for good health in quasi-military terms, to talk of “fighting” malaria or AIDS, to refer to a “campaign” of immunization or the “conquest” of smallpox, to “free” a population or a geographical area of some disease, to worry about the “arms race” that constantly occurs between pathogens and the drugs to hold them in check, to hope for a “silver bullet” against cancer or diabetes. In those terms, the providers of direct health services – whether aimed at individuals, communities or the environment – can be considered the front-line troops defending society against illness. But just as with an army, the health system must be much more than the soldiers in the field if it is to win any battles. Behind them is an entire apparatus to ensure that the fighters are adequately trained, informed, financed, supplied, inspired and led. It is also crucial to treat decently the population they are supposed to protect, to teach the “civilians” in the struggle how to defend themselves and their families, and to share equitably the burden of financing the war.

Unless those functions are properly carried out, firepower will be much less effective than it might be, and casualties will be higher. The emphasis here on overall results and on the functions more distant from the front line does not mean any denigration of the importance of disease control. It means rather to step back and consider what it is that the system as a whole is trying to do, and how well it is succeeding. Success means, among other things, more effective control of diseases, through better performance.

**Figure 2.1 Relations between functions and objectives of a health system**



## GOODNESS AND FAIRNESS: BOTH LEVEL AND DISTRIBUTION MATTER

A good health system, above all, contributes to good health. But it is not always satisfactory to protect or improve the average health of the population, if at the same time inequality worsens or remains high because the gain accrues disproportionately to those already enjoying better health. The health system also has the responsibility to try to reduce inequalities by preferentially improving the health of the worse-off, wherever these inequalities are caused by conditions amenable to intervention. The objective of good health is really twofold: the best attainable average level – *goodness* – and the smallest feasible differences among individuals and groups – *fairness*. A gain in either one of these, with no change in the other, constitutes an improvement, but the two may be in conflict. The logic is somewhat parallel to that concerning the distribution of income in a population. It is desirable to raise the average level, to reduce inequality, or both, and sometimes to judge the relative values of one and the other goal (with the difference that there is no argument for taking health away from anyone – health, unlike income or nonhuman assets, cannot be directly redistributed).

The distinction between the overall level and how it is distributed in the population also applies to responsiveness. Goodness means the system responds well on average to what people expect of it, with respect to its non-health aspects. Fairness means that it responds equally well to everyone, without discrimination or differences in how people are treated. The distribution of responsiveness matters, just as the distribution of health does. Either one is valuable by itself.

In contrast to the objectives of good health and responsiveness, there is no overall notion of goodness related to financing. There are good and bad ways to raise the resources for a health system, of course, but they are more or less good primarily as they affect how fairly the financial burden is shared. Fair financing, as the name suggests, is concerned only with distribution. It is not related to the total resource bill, nor to how the funds are used. While it is unambiguously preferable to have better health or a higher level of responsiveness, it is not always better to spend more on health because at high levels of expenditure there may be little additional health gain from more resources. The objectives of the health system do not include any particular level of total spending, either absolutely or relative to income. This is because, at all levels of spending, the resources devoted to health have competing uses, and it is a social choice – with no correct answer – how much to allocate to the health system. Nonetheless there is probably a minimum level of expenditure required to provide a whole population with a handful of the most cost-effective services, and many poor countries are currently spending too little even to assure that (4).

In countries where most health financing is private, and is largely out of pocket, no one makes this choice overall; it results from millions of individual decisions. As the level of prepayment rises, there are fewer and larger decisions, because spending is more and more determined by the policies and budgets of public entities and insurance funds. The public budget decision has the greatest effect in high income countries where most funding is government controlled or mandated, but in all countries it is one of the most basic public decisions. It is something that can be directly chosen, as the level of health outcome or of responsiveness cannot be.

## MEASURING GOAL ACHIEVEMENT

To assess a health system, one must measure five things: the overall level of health; the distribution of health in the population; the overall level of responsiveness; the distribution of responsiveness; and the distribution of financial contribution. For each one, WHO has used existing sources or newly generated data to calculate measures of attainment for the countries where information could be obtained. These data were also used to estimate values when particular numbers were judged unreliable, and to estimate attainment and performance for all other Member States. Several of these measures are novel and are explained in detail in the Statistical Annex, where all the estimates are given, along with intervals expressing the uncertainty or degree of confidence in the point estimate. The correct value for any indicator is estimated to have an 80% probability of falling within the uncertainty interval, with chances of 10% each of falling below the low value or above the high one. This recognition of inexactness underscores the importance of getting more and better data on all the basic indicators of population health, responsiveness and fairness in financial contribution, a task which forms part of WHO's continuing programme of work.

The achievements with respect to each objective are used to rank countries, as are the overall measures of achievement and performance described below. Since a given country or health system may have very different ranks on different attainments, Annex Table 1 shows the complete ranking for all Member States on all the measures. In several subsequent tables, countries are ranked in order of achievement or performance, and the order varies from one table to another. Since the ranking is based on estimates which include uncertainty as to the exact values, the rank assigned also includes uncertainty: a health system is not always assigned a specific position relative to all others but is estimated to lie somewhere within a narrower or broader range, depending on the uncertainties in the calculation. The ranks of different health systems therefore sometimes overlap to a greater or lesser degree, and two or more countries may have the same rank.

*Health* is the defining objective for the health system. This means making the health status of the entire population as good as possible over people's whole life cycle, taking account of both premature mortality and disability. Annex Table 2 presents three conventional and partial measures of health status, by country, without ranking: these are the probability of dying before age five years or between ages 15 and 59 years, and life expectancy at birth. For the first time, these measures are presented with estimates of uncertainty, and these uncertainties carry over to subsequent calculations. On the basis of the mortality figures, five strata are identified, ranging from low child and adult mortality to high child mortality and very high adult death rates. Combining these strata with the six WHO Regions gives 14 subregions defined geographically and epidemiologically (see the list of Member States by WHO Region and mortality stratum). Annex Table 3 presents estimates of mortality by cause and sex in 1999 in each of these subregions (not by country), and Annex Table 4 combines these death rates with information about disability to create estimates of one measure of overall population health: the burden of disease, that is, the numbers of disability-adjusted life years (DALYs) lost.

To assess overall population health and thus to judge how well the objective of good health is being achieved, WHO has chosen to use disability-adjusted life expectancy (DALE), which has the advantage of being directly comparable to life expectancy estimated from mortality alone and is readily compared across populations. Annex Table 5 provides estimates for all countries of disability-adjusted life expectancy. DALE is estimated to equal or exceed 70 years in 24 countries, and 60 years in over half the Member States of WHO. At

the other extreme are 32 countries where disability-adjusted life expectancy is estimated to be less than 40 years. Many of these are countries with major epidemics of HIV/AIDS, among other causes. Box 2.1 describes how these summary measures of population health are constructed and how they are related.

Figure 2.2 summarizes the relation between DALE and life expectancy without adjustment, for each of the 14 subregions, for both men and women. The adjustment is nearly uniform, at about seven years of healthy life equivalent lost to disability. Both absolutely and relatively this loss is slightly less for richer, low-mortality subregions, despite the fact that people live longer there and so have more opportunity to acquire non-fatal disabilities. Disability makes a substantial difference in poorer countries because some limitations – injury, blindness, paralysis and the debilitating effects of several tropical diseases such as malaria and shistosomiasis – strike children and young adults. Separating life expectancy into years in good health and years lived with disability therefore widens rather than narrows the difference in health status between richer and poorer populations. This is most evident in the share of life expectancy which is lost to disability: it ranges from less than 9% in the healthiest subregions to more than 14% in the least healthy. Annex Table 5 shows these shares for individual countries, where the range is even wider.

Annex Table 5 also provides estimates of health inequality. The distributional measure of health ranges from 1 for the case of perfect equality to zero for extreme inequality, which corresponds to a fraction of the population having an expectancy of 100 years and the rest

### Box 2.1 Summary measures of population health

No measure is perfect for the purpose of summing up the health of a population; each way of estimating it violates one or another desirable criterion. The two principal approaches are the burden of disease, which measures losses of good health compared to a long life free of disability, and some measure of life expectancy, adjusted to take account of time lived with a disability. Both ways of summarizing health use the same information about mortality and disability, and both are related to a survivorship curve, such as the bold line between the areas labelled *Disability* and *Mortality* in the figure.

The area labelled *Mortality* represents losses due to death, compared to a high standard of life expectancy: the burden of disease corresponds to all of that area plus a fraction of the area corresponding to time lived with disability. The fraction depends on the disability weights assigned to various states between death and perfect health. Life expectancy without any adjustment corresponds to the areas labelled *Survival free of disability* and *Disability* together, the whole area under the survivorship curve. Disability-adjusted life expectancy (DALE) then corresponds to the area for sur-

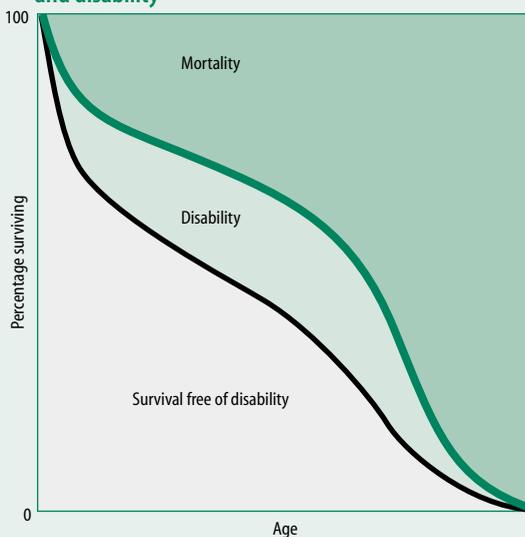
vival plus part of that for disability.

DALE is estimated from three kinds of information: the fraction of the population surviving to each age, calculated from birth and death rates; the prevalence of each type of disability at each age; and the weight assigned to each type of disability, which may or may not vary with age. Survival at each age is adjusted downward by the sum of all the disability

effects, each of which is the product of a weight and the complement of a prevalence (the share of the population not suffering that disability). These adjusted survival shares are then divided by the initial population, before any mortality occurred, to give the average number of equivalent healthy life years that a newborn member of the population could expect to live.

One important difference between the burden of disease estimation using disability-adjusted life years (DALYs) and that of DALE is that the former do, but the latter do not, distinguish the contribution of each disease to the overall result. DALE has the advantage that it does not require as many choices of parameters for the calculation, and it is directly comparable to the more familiar notion of life expectancy without adjustment.

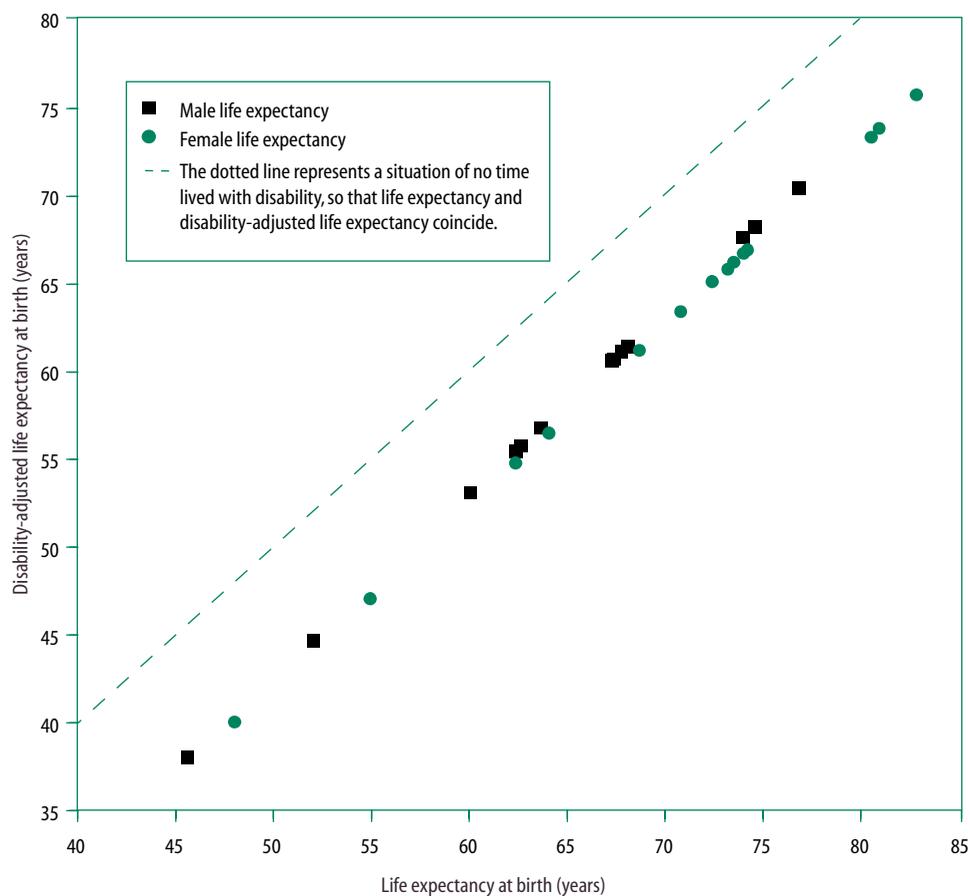
Summarizing population health from mortality and disability



Source: Murray CJL, Salomon JA, Mathers C. *A critical examination of summary measures of population health*. Geneva, World Health Organization, 1999 (GPE Discussion paper No. 12).

having no expectation of surviving infancy. *If everyone had the same life expectancy, adjusted for disability, the system would be perfectly fair with respect to health, even though people would actually die at different ages.* For a small number of countries it has been possible to estimate the distribution of life expectancy within the population using information on both child

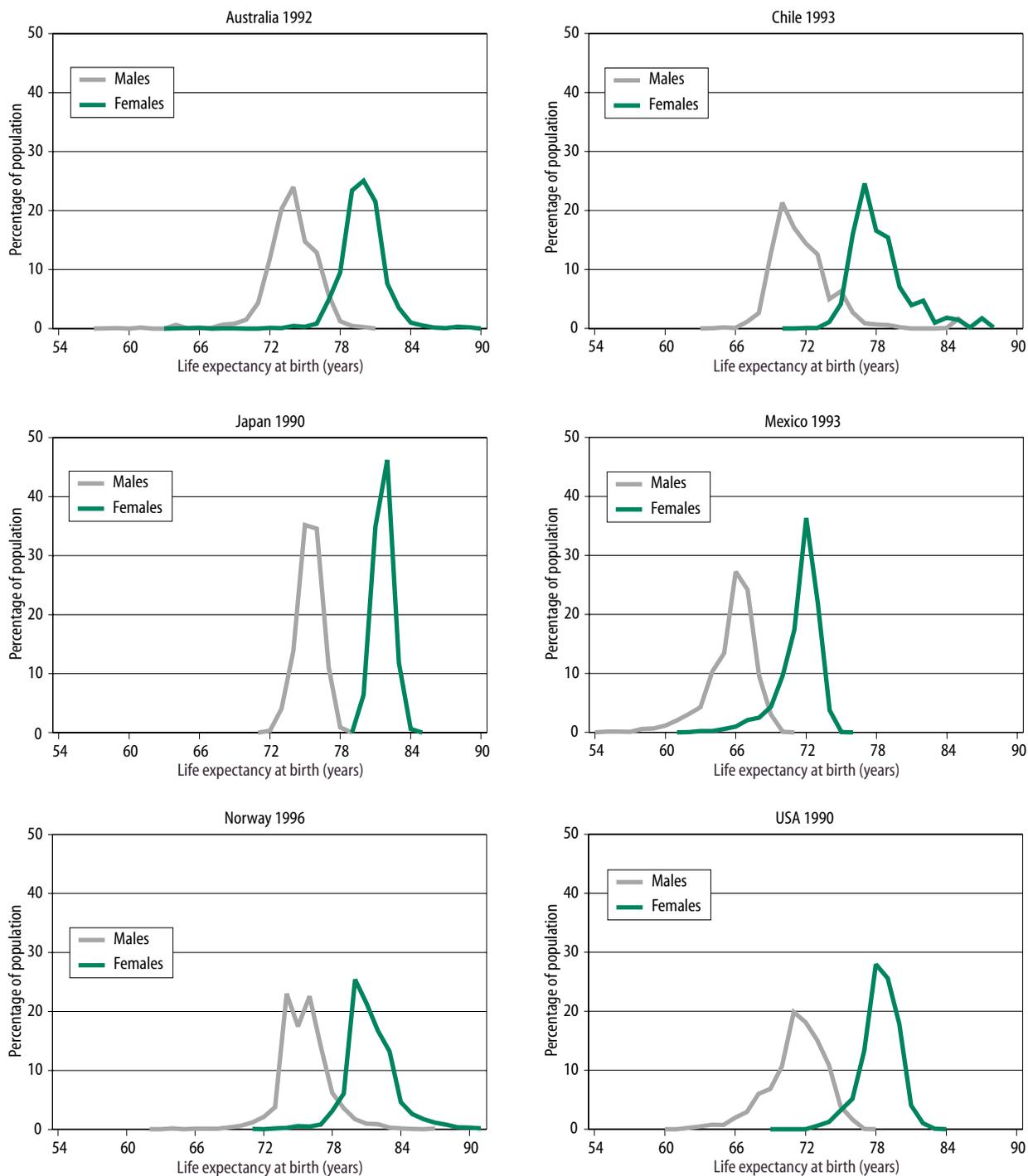
**Figure 2.2 Life expectancy and disability-adjusted life expectancy for males and females, by WHO Region and stratum defined by child mortality and adult mortality, 1999**



WHO Region	Mortality stratum		Males		Females	
	Child	Adult	Life expectancy	Disability adjusted	Life expectancy	Disability adjusted
AFR	High	High	52.0	44.6	54.9	47.0
	High	Very high	45.6	38.0	48.0	40.0
AMR	Very low	Very low	73.9	67.5	80.4	73.2
	Low	Low	67.3	60.6	74.1	66.8
	High	High	63.6	56.7	68.6	61.1
EMR	Low	Low	67.7	61.0	70.7	63.3
	High	High	60.0	53.0	62.3	54.7
EUR	Very low	Very low	74.5	68.1	80.8	73.7
	Low	Low	67.3	60.6	73.9	66.6
	Low	High	62.3	55.4	73.4	66.1
SEAR	Low	Low	67.2	60.5	73.1	65.7
	High	High	62.6	55.7	64.0	56.4
WPR	Very low	Very low	76.7	70.3	82.7	75.6
	Low	Low	68.0	61.3	72.3	65.0

and adult mortality; these results are presented below. For most countries, however, it has so far been possible to use only child mortality data. Because high-income countries have largely eliminated child mortality, the highest ranking countries in Annex Table 5 nearly all have relatively high incomes; most are European. A few Latin American countries which

Figure 2.3 Inequality in life expectancy at birth, by sex, in six countries



have made great progress in controlling child deaths also show considerable equality of health. Except for Afghanistan and Pakistan, all the countries ranked lowest on child health equality are in sub-Saharan Africa, where child mortality is still relatively high. When more complete data are available on inequalities in adult mortality they will be used in future WHO estimates, and these rankings will change, because high income countries differ more in adult than in child death rates.

Inequalities in life expectancy persist, and are strongly associated with socioeconomic class, even in countries with quite good health status on average (5). Figure 2.3 illustrates these inequalities for six countries, showing the distribution of life expectancy at birth for both men and women, using data on adult as well as child mortality, estimated from large numbers of small-area studies which cover the entire country. Among these six countries, health is most equally distributed in Japan. Both distributions of life expectancy are sharply peaked, concentrating the whole population of either men or women in a range of only about six years. There is far more inequality in Mexico and in the United States, and in both cases that arises because part of the population has a much lower expectation than the rest, after age five years. The inequality is particularly marked for men. An opposite pattern characterizes Chile, which shows very high equality in child health: the degree of adult inequality is about the same as for Mexico and the United States, but it arises because part of the population has an unusually high life expectancy. Australia and Norway both show more symmetric distributions. These results emphasize the value of judging health system achievement not only by averages or overall levels but by seeing whether everyone has about the same expectation of life.

*Responsiveness* is not a measure of how the system responds to health needs, which shows up in health outcomes, but of how the system performs relative to non-health aspects, meeting or not meeting a population's expectations of how it should be treated by providers of prevention, care or non-personal services. (The last category is least important, since individuals normally do not come into personal contact with such interventions. However, even public health measures such as vector control can be conducted with more or less respect for people and their wishes. Assessing the responsiveness of providers of non-personal services is a particular challenge.)

Some systems are highly unresponsive. The Soviet health system prior to 1990 had become highly impersonal and inhuman in the way it processed people. A common complaint in many countries about public sector health workers focuses on their rudeness and arrogance in relations with patients (6, 7). Waiting times for non-emergency surgery vary considerably among industrialized countries (8) and are the subject of much criticism of ministries of health (9). Recognizing responsiveness as an intrinsic goal of health systems establishes that these systems are there to serve people, and involves more than an assessment of people's satisfaction with the purely medical care they receive.

The general notion of responsiveness can be decomposed in many ways. One basic distinction is between elements related to respect for human beings as persons – which are largely subjective and judged primarily by the patient – and more objective elements related to how a system meets certain commonly expressed concerns of patients and their families as clients of health systems, some of which can be directly observed at health facilities. Subdividing these two categories leads to seven distinct elements or aspects of responsiveness.

Respect for persons includes:

- Respect for the dignity of the person. At the extreme, this means not sterilizing individuals with a genetic disorder or locking up people with communicable diseases, which would violate basic human rights. More generally, it means not humiliating or demeaning patients.
- Confidentiality, or the right to determine who has access to one's personal health information.
- Autonomy to participate in choices about one's own health. This includes helping choose what treatment to receive or not to receive.

Client orientation includes:

- Prompt attention: immediate attention in emergencies, and reasonable waiting times for non-emergencies.
- Amenities of adequate quality, such as cleanliness, space, and hospital food.
- Access to social support networks – family and friends – for people receiving care.
- Choice of provider, or freedom to select which individual or organization delivers one's care.

In general, responsiveness contributes to health by promoting utilization, but that is not always the case. Greater autonomy can mean that people do not take up an intervention because they perceive the individual benefit to be small or the risk to be substantial, and do not value the collective or population benefit. This is particularly likely for immunization, especially if there is fear of adverse reactions. Individual freedom to choose whether or not to be immunized is in conflict with the public health objective of high coverage to prevent epidemics. Such conflict has occurred, for example, in the United Kingdom for pertussis and in Greece for rubella vaccine (10). The overall performance of a health system may therefore involve trade-offs among objectives.

Opinions on how well a health system performs on such subjective dimensions as responsiveness might be influenced by any of a number of features of the systems themselves, or of the respondents. Since poor people may expect less than rich people, and be more satisfied with unresponsive services, measures of responsiveness should correct for

### Box 2.2 How important are the different elements of responsiveness ?

The key informant survey, consisting of 1791 interviews in 35 countries, yielded scores (from 0 to 10) on each element of responsiveness, as well as overall scores. A second, Internet-based survey of 1006 participants (half from within WHO) generated opinions about the relative importance of the elements, which were used to combine the element scores into an overall score instead of just taking the mean or using the key informants' overall responses.

Respondents were asked to rank the seven elements in order of importance, and the weights were derived from the frequencies with which an element was ranked first, second, and so on. Respect for persons and client orientation were rated as equally important overall, and the three elements of respect for persons were also regarded as all about equally important. The four elements of

client orientation received different rankings and therefore unequal weights. The final weights are shown in the table.

Analysis of the element scores themselves, as estimated by the key informants, showed three consistent biases: for the same country, women respondents gave lower scores than men, and government officials gave higher scores than more independent informants; and all informants' scores tended to be higher for countries with less political freedom, as measured by a composite index. The data were adjusted to make the scores comparable across countries by removing the influence of these factors, so that all the scores are estimates of the ratings that would be given in a politically free country, by respondents who did not work for the government, half of whom were women.

<i>Respect for persons</i>	
Total	50%
Respect for dignity	16.7%
Confidentiality	16.7%
Autonomy	16.7%
<i>Client orientation</i>	
Total	50%
Prompt attention	20%
Quality of amenities	15%
Access to social support networks	10%
Choice of provider	5%

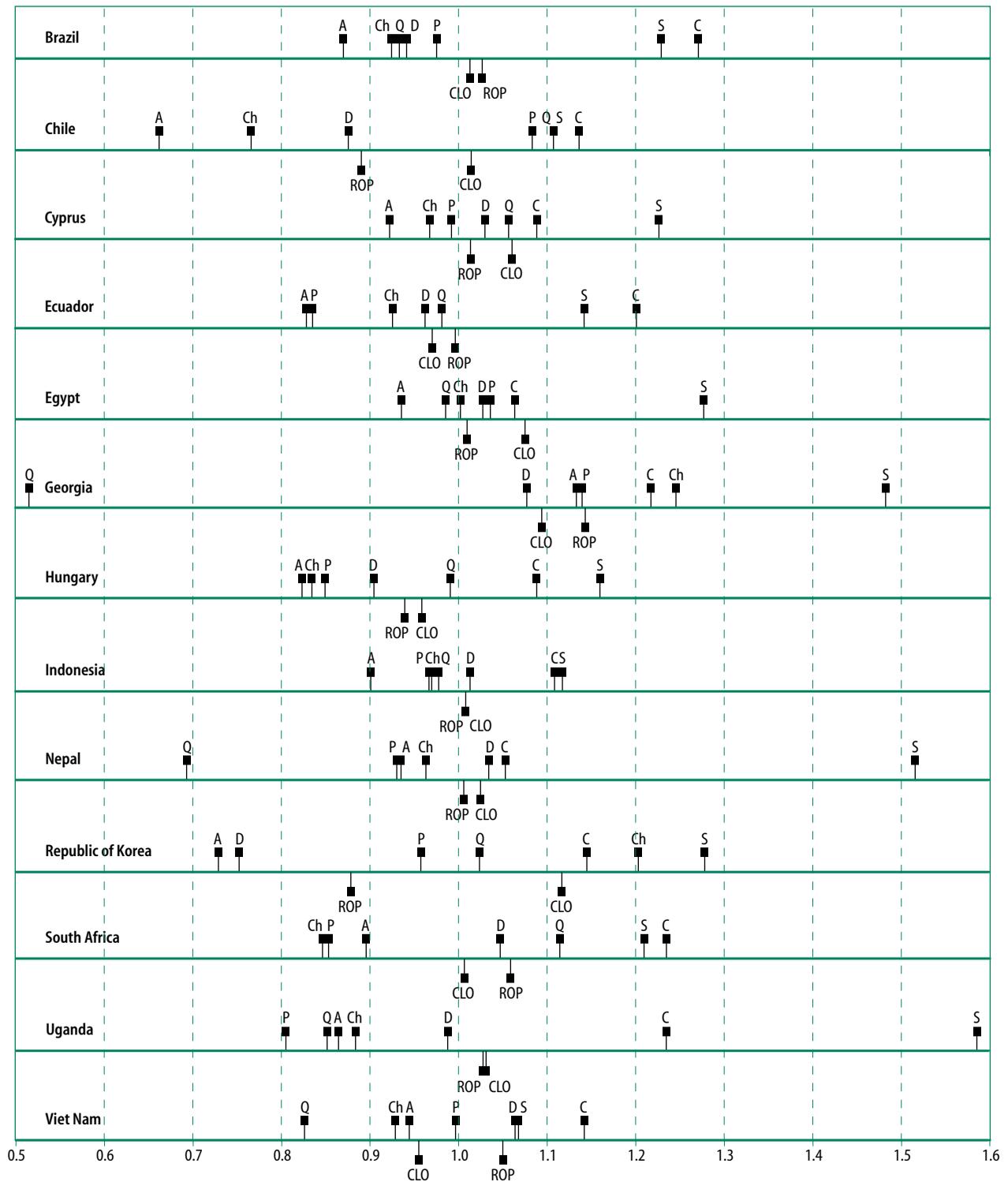
these differences, as well as for cultural differences among countries (11). Even without such adjustment, comparisons of how knowledgeable observers rate health system achievements can reveal on which aspects of responsiveness a system seems to satisfy its users best. Judgements about average level and inequality of the components of responsiveness were developed in each of 35 countries by a network of 50 or more key informants. A separate survey of over a thousand respondents was used to develop weights for combining these scores into an overall rating. Box 2.2 describes the results of this exercise. Estimates for other Member States were derived from the 35 observations, adjusted for differences among countries and informant groups. Surveys of population opinion and direct observation of health provision can both be used to complement these judgements.

Figure 2.4 illustrates in detail the scores of the seven individual elements, relative to the overall score, within each of 13 countries chosen to reflect all WHO Regions and typical of the entire set of countries studied. The health systems examined always appear to perform relatively well on the two dimensions of access to social support networks and confidentiality, sometimes very much better than on other aspects. The systematically high rating for social support may reflect a trade-off against the quality of amenities, because a health care facility that cannot, for lack of resources, offer good quality food or non-medical attention can compensate for that by allowing relatives and friends to attend to patients' needs. One reason why confidentiality seems not to be a problem in these countries may be that there is little private insurance and therefore little risk of coverage being denied because a provider reveals some information about a patient. There is somewhat less consistency at the other end of the scale, but autonomy is among the three lowest-rated elements of responsiveness 34 times out of 35 – and the lowest ranked element almost half the time – and performance is also often poor with respect to choice of provider and promptness of care.

As with health status, it is not only overall responsiveness that matters, if some people are treated with courtesy while others are humiliated or disdained. *A perfectly fair health system would make no such distinctions, and would receive the same rating of responsiveness on every element, for every group in the population.* In almost every country where key informants were surveyed, the poor were identified as the main disadvantaged group. In particular, they were considered to be treated with less respect for their dignity, to have less choice of providers and to be offered poorer quality amenities than the non-poor. In nearly as many cases, rural populations – among whom the poor are concentrated – were regarded as being treated worse than urban dwellers, suffering especially from less prompt attention, less choice of providers and lower quality of amenities. Some respondents in one or several countries also identified women, children or adolescents, indigenous or tribal groups or others as receiving worse treatment than the rest of the population.

The elements of client orientation, where the poor and the rural population are less well treated, all have economic implications: it generally costs more to assure quick attention and to offer high quality food, more space and well-kept facilities. It also makes cost control harder if people are allowed to choose their providers, and costs differ among them. The strongest associations occur for quality of basic amenities and promptness of attention. The former is closely related to income per head and to the share of private expenditure in total health spending; the latter is closely related to average years of schooling of the population, which is also associated with income. In contrast, the elements of respect for persons can be costless, apart perhaps from some training of providers and administrators. These elements – respect for dignity, autonomy, and confidentiality – show no relation to health system spending. There is scope for improving health system performance in these respects without taking any resources away from the primary objective of better health. This

Figure 2.4 Relative scores of health system responsiveness elements, in 13 countries, 1999



All scores are normalized relative to the average overall country responsiveness score = 1.

A = autonomy, C = confidentiality, Ch = choice of provider or facility, D = dignity, P = promptness, Q = quality of basic amenities, S = access to social support networks, ROP = respect of persons, CLO = client orientation.

is particularly the case for autonomy, where performance is often poorly rated.

Annex Table 6 reports adjusted scores for overall responsiveness, as well as a measure of fairness based on the informants' views as to which groups are most often discriminated against in a country's population and on how large those groups are. Either a larger group being affected, or more informants agreeing on that group's being treated worse than some others, implies more inequality of responsiveness and therefore less achievement of fairness. Since some elements of responsiveness are costly, it is not surprising that most of the highest ranked countries spend relatively large amounts on health. They are also often countries where a large share of provision is private, even if much of the financing for it is public or publicly mandated. However, the association with a country's income or health expenditure is less marked than it is for health status. Several poor African and Asian countries rank fairly high on the level of responsiveness. And countries that perform well on average for responding to people's expectations may nonetheless rank much lower on the distributional index.

*Fair financing* in health systems means that the risks each household faces due to the costs of the health system are distributed according to ability to pay rather than to the risk of illness: a fairly financed system ensures financial protection for everyone. A health system in which individuals or households are sometimes forced into poverty through their purchase of needed care, or forced to do without it because of the cost, is unfair. This situation characterizes most poorer countries and some middle and high income ones, in which at least part of the population is inadequately protected from financial risks (12).

Paying for health care can be unfair in two different ways. It can expose families to large *unexpected* expenses, that is, costs that could not be foreseen and have to be paid out of pocket at the moment of utilization of services rather than being covered by some kind of prepayment. Or it can impose *regressive* payments, in which those least able to contribute pay proportionately more than the better-off. The first problem is solved by minimizing the share of out-of-pocket financing of the system, so as to rely as fully as possible on more predictable prepayment that is unrelated to illness or utilization. The second is solved by assuring that each form of prepayment – through taxes of all kinds, social insurance, or voluntary insurance – is progressive or at least neutral with respect to income, being related to capacity to pay rather than to health risk.

Out-of-pocket payments are generally regressive but they can, in principle, be neutral or progressive. When this happens, and out-of-pocket expenses are not too large, they need not impoverish anyone or deter the poor from obtaining care. However, of all the forms of financing they are the most difficult to make progressive. Arrangements that exempt the destitute from user fees at public facilities, or impose a sliding scale based on socioeconomic characteristics, are attempts to reduce the risk associated with out-of-pocket payments (13, 14). Except when private practitioners know their clientele well enough to discriminate among them in fees – and the better-off accept that their charges will subsidize the worse-off – such arrangements are limited to public facilities, which often account for only a small share of utilization in poor countries. And even then, such schemes require relatively high administrative costs to distinguish among users, and typically affect only a small amount of total risk-related payments.

For this reason, financial fairness is best served by more, as well as by more progressive, prepayment in place of out-of-pocket expenditure. And the latter should be small not only in the aggregate, but relative to households' ability to pay. Prepayment that is closely related to *ex ante* risk, as judged from observable characteristics – risk-related insurance premiums,

for example – is still preferable to out-of-pocket payment because it is more predictable, and may be justified to the extent that the risks are under a person's control. However, the ideal is largely to disconnect a household's financial contribution to the health system from its health risks, and separate it almost entirely from the use of needed services. The question of how far insurance prepayments may be related to risks, and how such premiums should be financed, including subsidies for those unable to pay, is treated in Chapter 5.

*Ex post*, the burden of health financing on a particular household is the share that its actual health expenses are of its capacity to pay. The numerator includes all costs attributable to the household, including those it is not even aware of paying, such as the share of sales or value-added taxes it pays on consumption, which governments then devote to health, and the contribution via insurance provided, and partly financed, by employers.

The denominator is a measure of the household's capacity to pay. In poor households, a large share goes for basic necessities, particularly food, whereas richer households have more margin for other spending, including spending on health care. Food spending is treated as an approximation to expenditure on basic needs. Total non-food spending is taken as an approximation of the household's discretionary and relatively permanent income, which is less volatile than recorded income (15) and a better measure of what a household can afford to spend on health and other non-food needs.

In sum, *the way health care is financed is perfectly fair if the ratio of total health contribution to total non-food spending is identical for all households, independently of their income, their health status or their use of the health system.* This indicator expresses the trenchant view of Aneurin Bevan, that "The essence of a satisfactory health service is that the rich and the poor are treated alike, that poverty is not a disability, and wealth is not advantaged." (16). Clearly the financing would be unfair if poor households spent a larger share than rich ones, either because they were less protected by prepayment systems and so had to pay relatively more out of pocket, or because the prepayment arrangements were regressive. But to identify fairness with equality means that the system is also regarded as unfair if *rich* households pay more, as a share of their capacity. Simply by paying the same fraction as poor households, they would be subsidizing those with lower capacity to pay. It is true that well-off households might choose to pay still more, particularly by buying more insurance, but that can be considered equitable only if the extra spending is prepaid and if the choice is entirely voluntary and not determined by the system of taxes or mandatory insurance contributions.

Families that spend 50% or more of their non-food expenditure on health are likely to be impoverished as a result. Detailed household surveys show that in Brazil, Bulgaria, Jamaica, Kyrgyzstan, Mexico, Nepal, Nicaragua, Paraguay, Peru, the Russian Federation, Viet Nam and Zambia more than 1% of all households had to spend on health half or more of their full monthly capacity to pay, which means that in large countries millions of families are at risk of impoverishment. Invariably the reason is high out-of-pocket spending. This high potential for financial catastrophe has much to do with how the health system is financed, and not only with the overall level of spending or the income of the country.

The fairness of the distribution of financial contribution is summarized in an index which is inversely related to the inequality in the distribution, and presented in Annex Table 7. The index runs from zero (extreme inequality) to 1 (perfect equality). For most countries, and particularly for most high income countries, the value is not far from 1, but great inequality characterizes a few countries in which nearly all health spending is out-of-pocket, notably China, Nepal and Viet Nam. However, in some countries where most spending is out-of-

**Figure 2.5 Household contributions to financing health, as percentage of capacity to pay, in eight countries**



pocket, there is nonetheless little inequality because that spending is relatively progressive and few families spend as much as half their non-food expenditure on health. Bangladesh and India are examples. Generally, high values of equality are associated with predominantly prepaid financing, but Brazil shows extreme inequality despite a high share of prepayment, because of the great inequality in incomes and the large number of families at risk of impoverishment.

The summary measure of fairness does not distinguish poor from rich households. Figure 2.5 introduces this distinction, by showing how the burden is distributed across deciles of capacity to pay, and divided between prepayment and out-of-pocket spending, in eight low and middle income countries. Prepayment is clearly progressive – the rich contribute a larger share – in Mexico and the United Republic of Tanzania, and also in Bangladesh and Colombia (not shown). It is actually regressive in India and Pakistan, and also in Guyana, Kyrgyzstan, Nepal, Peru and the Russian Federation (not shown). In other countries – Brazil, Bulgaria, Jamaica, Nicaragua, Paraguay, Romania and Zambia – the prepaid contribution is distributed more or less neutrally or varies irregularly. Out-of-pocket spending shows more variation, as might be expected; for example, it is progressive in India and quite regressive in Pakistan and Viet Nam, where there is almost no prepaid financing at all.

Total non-food spending also includes whatever the household spends out of pocket on health care. That spending is largely unpredictable or transitory, so to include it may overstate the family's capacity to pay. If out-of-pocket expenditure is small, it makes no difference; but if it is large, it may have been financed by selling assets, going into debt, requiring more family members to work or for some to take on more hours or other employments, or even temporarily reducing consumption of necessities. If household capacity to pay is defined as non-food spending *less* out-of-pocket health spending, then families with large out-of-pocket expenditures are classified as poor, instead of being scattered throughout the population. The way the health system is financed then looks systematically less fair in most countries, and the culprit is always the large share of out-of-pocket spending. Pre-

### Box 2.3 What does fair contribution measure and not measure ?

The way fair contribution to health care finance is measured is strictly *ex post*, referring to what households actually contribute rather than to their *ex ante* risks of needing health care. That means that there is no need to estimate the "coverage" of the population by different risk-sharing schemes. Coverage in financial rather than nominal terms – how much people are really protected, not simply whether they have insurance or participate in social security – is hard to estimate beforehand, and in any case such coverage is, like accessibility, an instrumental rather than a final goal. Nominal coverage does provide people with a sense of security which also

affects their spending and saving decisions, but that is not a goal in itself.

People who do not use care when they need it, because they cannot afford the out-of-pocket cost, appear to spend less than they really need to. Estimating what they *would* have spent if they could afford it would give a different distribution of contributions, and would almost surely show even more unfairness. Basing the measure on what is actually spent – which is all that the data allow – overstates the degree to which a health system achieves a fair distribution of the financing burden.

The measure also says nothing about how a family obtains the cash

to pay out of pocket for health care (or for some forms of prepayment such as "health cards" or vouchers). Households much of whose income is in kind rather than cash may forego health care because they cannot obtain the cash when needed, and the data will show only that they did not spend. Without further analysis there is no way to distinguish illiquidity from all the other reasons why a health need did not eventuate in expenditure. A less serious but sometimes still consequential liquidity problem arises when a household has to pay out of pocket for care, and then wait for reimbursement from an insurer. This need to finance care temporarily arises for populations wealthy

enough to have formal – usually private – insurance.

Finally, and most important, fair financing means only equity in how the financial burden of supporting a health system is shared. It says nothing about whether the *utilization* of health services is fair, which is an equally crucial issue in the overall fairness of the system. Fair financing is concerned with the principle of *from each according to ability*, but not with the principle of *to each according to need*. Unfairness in use relative to need shows up in inequalities in health status, because service utilization ought to reduce such inequalities so far as they are amenable to intervention.

payment in low income countries is commonly too small a share of the total to offset the regressive and very unequal impact of out-of-pocket spending. What is worse, in many countries there is *no* offsetting effect because prepayment via taxes is also regressive. In those cases the poor thus suffer twice – all of them have to pay an unfair share whether or not they use health services, and then some of them have also to pay an even more unfair contribution out of pocket. These are the strongest findings to emerge from the analysis of financing; their implications are developed further in Chapter 5. Box 2.3 discusses four other features of this way of measuring fairness in financing.

Much of the analytical effort behind this report went into developing more and better information about expenditure on health and constructing national health accounts. As described further in Chapter 6, knowledge of where resources are coming from, through what channels they flow and how they are used, is crucial to better stewardship of the system. Annex Table 8 presents the estimates of total health spending, its separation into private and public sources, the distinction between tax-financed and social security health spending in the public sector, and that between insurance and out-of-pocket spending in the private sector, and the overall distinction between prepayment and out-of-pocket spending that helps determine how fairly health systems are paid for. These data, besides being of direct interest, have been used to check the estimates of household expenditure discussed above and to estimate values for indicators that are strongly related to spending.

Whatever the sources and distribution of finance, the level of resources devoted to health is an input into the system, not an outcome: it is what makes the outcomes possible, and against which the system's achievements should be evaluated. The next two sections take up the question of how best to do this, first by developing an overall measure of attainment and then by relating that achievement to resource use, as a measure of performance.

#### Box 2.4 Weighting the achievements that go into overall attainment

To derive a set of weights for the different achievements that compose overall attainment, WHO conducted a survey of 1006 respondents from 125 countries, half from among its own staff. The questions were designed to elicit not only views about how important each goal is relative to the others (for example, responsiveness compared to health status), but also opinions about what kind of inequality matters most. The responses were checked for consistency and bias, and yielded nearly identical values in each of many different groups – poorer versus richer countries, men versus women, WHO staff versus other respondents. The final weights are shown in the table.

As expected, health is regarded as the most important of the objectives, clearly the primary or defining goal of a system. But fully half of the concern for health is a concern for equality, not simply for a high average. Taking "health" apart into two goals emphasizes the great value of fairness, and not only of goodness. This is fully consistent with WHO's concentration on the poor, the least healthy, the worst-off in society. Equal weights also result from the survey for the overall level and for distribution or equality where responsiveness is concerned. In total, how the system treats people in non-health aspects is as important as either health level or health

equality. And fairness in how health is paid for, which is not a major traditional concern of WHO or the ministries of health it deals with and supports, receives the relatively large weight of one-fourth, equal to that for responsiveness. Both in this case and in that of responsiveness, the weight assigned by respondents probably reflects the direct or intrinsic importance of the objective, and also the indirect or instrumental contribution it makes to achieving good health; it is difficult to separate these two aspects. There is clear agreement that a well-functioning health system should do much more than simply promote the best possible level of overall health.

The exercise of weighting the five objectives also provides values for the relative importance of goodness and fairness. Together, the *levels* of health and of responsiveness receive a weight of three-eighths of the total. The three distributional measures, which together describe the *equity* of the system, account for the remaining five-eighths. Countries which have achieved only rather short life expectancies and cannot adequately meet their peoples' expectations for prompt attention or amenities may nonetheless be regarded as having health systems which perform well with respect to fairness on one or more dimensions.

<i>Health (disability-adjusted life expectancy)</i>	
Total	50%
Overall or average	25%
Distribution or equality	25%
<i>Responsiveness</i>	
Total	25%
Overall or average	12.5%
Distribution or equality	12.5%
<i>Fair financial contribution</i>	
Distribution or equality	25%

## OVERALL ATTAINMENT: GOODNESS AND FAIRNESS COMBINED

To the extent that a health system achieves a long disability-adjusted life expectancy, or a high level of responsiveness (or a high degree of equality in either or both), or a fair distribution of the financing burden, it can be said to perform well with respect to that objective. Since a system can do well on one or more dimensions and poorly on others, comparison across countries or through time requires that the five goals be summed into a single overall measure. There is no natural scale on which to add together years of life, responsiveness scores, and measures of inequality or fairness, so combining the measures of achievement means assigning a weight or relative importance to each one. Box 2.4 describes the procedure and the results.

Applying these weights to the achievements described in Annex Tables 5, 6 and 7 yields an overall attainment score for each health system. These scores are presented in Annex Table 9, together with an estimate of the uncertainty around each value, derived from the uncertainties for the components. Because rich countries generally enjoy good health, and because high incomes allow for large health expenditures which are also predominantly prepaid and often largely public, the ranking by overall attainment is closely related to income and health spending. However, the large weight given to distributional goals explains why, for example, Japan outranks the United States and why Chile, Colombia and Cuba outrank all other Latin American countries. It is not surprising that, with three Asian exceptions, the 30 worst-off countries are all in Africa.

## PERFORMANCE: GETTING RESULTS FROM RESOURCES

The overall indicator of attainment, like the five specific achievements which compose it, is an absolute measure. It says how well a country has done in reaching the different goals, but it says nothing about how that outcome compares to what might have been achieved with the resources available in the country. It is *achievement relative to resources* that is the critical measure of a health system's performance.

Thus if Sweden enjoys better health than Uganda – life expectancy is almost exactly twice as long – that is in large part because it spends exactly 35 times as much per capita on its health system. But Pakistan spends almost precisely the same amount per person as Uganda, out of an income per person that is close to Uganda's, and yet it has a life expectancy almost 25 years higher. This is the crucial comparison: why are health outcomes in Pakistan so much better, for the same expenditure? And it is health expenditure that matters, not the country's total income, because one society may choose to spend less of a given income on health than another. Each health system should be judged according to the resources actually at its disposal, not according to other resources which in principle could have been devoted to health but were used for something else.

Health outcomes have often been assessed in relation to inputs such as the number of doctors or hospital beds per unit of population. This approach indicates what these inputs *actually* produce, but it tells little about the health system's *potential* – what it could do if it used the same level of financial resources to produce and deploy different numbers and combinations of professionals, buildings, equipment and consumables. In these comparisons, the right measure of resources is money, since that is used to buy all the real inputs.

To assess relative performance requires a scale, one end of which establishes an upper

limit or “frontier”, corresponding to *the most that could be expected of a health system*. This frontier – derived using information from many countries but with a specific value for each country – represents the level of attainment which a health system might achieve, but which no country surpasses. At the other extreme, a lower boundary needs to be defined for *the least that could be demanded of the health system* (17). With this scale it is possible to see how much of this potential has been realized. In other words, comparing actual attainment with potential shows how far from its own frontier of maximal performance is each country’s health system.

WHO has estimated two relations between outcomes and health system resources. One estimate relates resources only to average health status (disability-adjusted life expectancy, DALE), which makes it somewhat comparable to many previous analyses of performance in health. The other relates resources to the overall attainment measure based on all five objectives. The same value of total resources is used for a country in both cases, because there is no way to identify expenditure as being directed to producing health services, determining responsiveness or making the financing more or less fair. The same is true of resources used to improve the distribution of health or responsiveness, rather than the average level.

Each frontier is a function of one other variable besides health system expenditure. That is the average years of schooling in the adult population, which is a measure of human

### Box 2.5 Estimating the best to be expected and the least to be demanded

WHO’s estimates of the upper and lower bounds of health system performance differ in two important ways from most analyses of what health systems actually achieve. The first is that a “frontier” is meaningful only if no country can lie beyond it, although at least one must lie on it. The frontier or upper limit is therefore estimated by a statistical technique which allows for errors in one direction only, minimizing the distances between the frontier and the calculated performance values. (The lower bound is estimated by the conventional technique of allowing errors in either direction.) The second is that the object is not to *explain* what each country or health system has attained, so much as to form an estimate of what should be possible. The degree of explanation could be increased by introducing many more variables. If tropical countries show systematically lower achievement in health, because of the effects of many diseases concentrated near the equator, a variable indicating

tropical location would raise the explanatory or predictive power. Similarly, if outcomes are worse with respect to equality in ethnically diverse countries, a variable reflecting that heterogeneity would explain the outcomes observed.

The difficulty with the attempt to explain as much as possible is that it leads to a different frontier, according to every additional variable. There would be one for tropical countries and another for colder climates; one for ethnically mixed countries and another for those with more uniform populations; and so on. If performance were measured relative to the frontier for each type of country, almost every health system might look about equally efficient in the use of resources, because less would be expected of some than of others. Every additional explanation would be the equivalent of a reason for not doing better. This is particularly true of explanations related to individual diseases: AIDS and malaria are major causes of health loss in many sub-Saharan African countries, but to include

their effects in the estimation of the frontier means judging those countries only according to how well they control all *other* diseases, as though nothing could be done about AIDS and malaria. This is the reason for estimating the frontier according to nothing but expenditure and human capital, which is a general measure of society’s capacity for many kinds of performance, including performance of the health system.

The measures of attainment draw on data referring to the past several years, to make the estimates more robust and less susceptible to anomalous values in any one year. The measures of expenditure and human capital are similarly constructed from more than one year’s data. Nonetheless, both the outcomes and the factors that determine potential performance are meant to describe the current situation of countries. They do not take into account how past decisions and use of resources may have limited what a system can actually achieve today – which could also be a rea-

son for poor performance – nor do they say how quickly a poorly performing system might be expected to improve and come closer to the frontier.

This way of estimating what is feasible bypasses two particularly complex issues which are well illustrated by control of tobacco-related mortality and disability. One is that many actions taken by health systems produce results only after a number of years, so that resources used today are not closely related to outcomes today. If a health system somehow persuaded all smokers to quit and no one to take up the habit, it would be many years before there was no more tobacco-induced disease burden.<sup>1</sup> The other is that no health system could reasonably be expected to bring smoking prevalence down to zero any time soon, no matter how hard it tried. Determining how to evaluate progress rather than only a health system’s current performance is one of many challenges for future effort.

<sup>1</sup> Jha P, Chaloupka F, eds. *Tobacco control policies in developing countries*. Oxford, Oxford University Press for the World Bank and the World Health Organization, 2000.

capital and therefore of the long-run potential, if not the current or actual, state of development of the country. It is a proxy for most of the factors outside the health system that contribute to health status, and probably also to the degree of responsiveness and to how health is financed. Box 2.5 explains how the upper and lower limits are estimated and how they should be interpreted.

Since the estimation is based entirely on country data rather than a model of what is ideal or feasible, and since there are upper limits to all the achievements, the frontier rises rapidly with additional resources when spending is low, and then rises more and more slowly as expenditure reaches the levels typical of rich countries. A health system can move *towards* the frontier by improving performance, that is by achieving more with the same resources. It can move *along* the frontier by spending more or less on health and reaching a different level of attainment but the same degree of performance. The entire frontier can also move *outward*, as new knowledge makes it possible to achieve better health or other outcomes, for given health system resources and a given level of human capital. Most of the enormous improvement in health over the last century and a half, described in Chapter 1, is due to such an expansion or outward movement of what it is possible to achieve.

If there were no health system in the modern sense, people would still be born, live and die; life expectancy would be much less than now, but it would not be zero. There would be no expenditure on health and hence no question of how fairly the financial burden was distributed. Similarly, there would be no responsiveness. So the minimum level of achievement would involve only health status, and in the absence of information about inequalities, only the average level of health. In the measure of overall attainment the values for the other four objectives, including all those related to inequality, would be set at zero. To estimate this minimum, WHO has used information from a limited number of countries *circa* 1900, relating life expectancy – with no adjustment for disability – to estimates of income. The situation at the turn of the last century is taken as the starting point for the great advances made possible by increased knowledge, investment and resources devoted to health. Some of the changes have the effect of raising the minimum – the eradication of smallpox is the best example. The emergence of HIV/AIDS and of tobacco-related disease have the opposite effect, making it harder than it was in 1900 to achieve a given level of health.

The question for any health system today is, given the country's human capital and the resources devoted to its health system, how close has it come to the most that could be asked of it? Relating outcomes in this way to the estimated minimum and maximum attainments and to the use of economic resources defines the overall indicator of system *performance*: to perform well means to move away from the minimum attainment and come close to the maximum. In economic terms, performance is a measure of efficiency: an efficient health system achieves much, relative to the resources at its disposal. In contrast, an inefficient system is wasteful of resources, even if it achieves high levels of health, responsiveness and fairness. That is, it could be expected to do still better, because countries spending less do comparably well or countries spending a little more achieve much better outcomes.

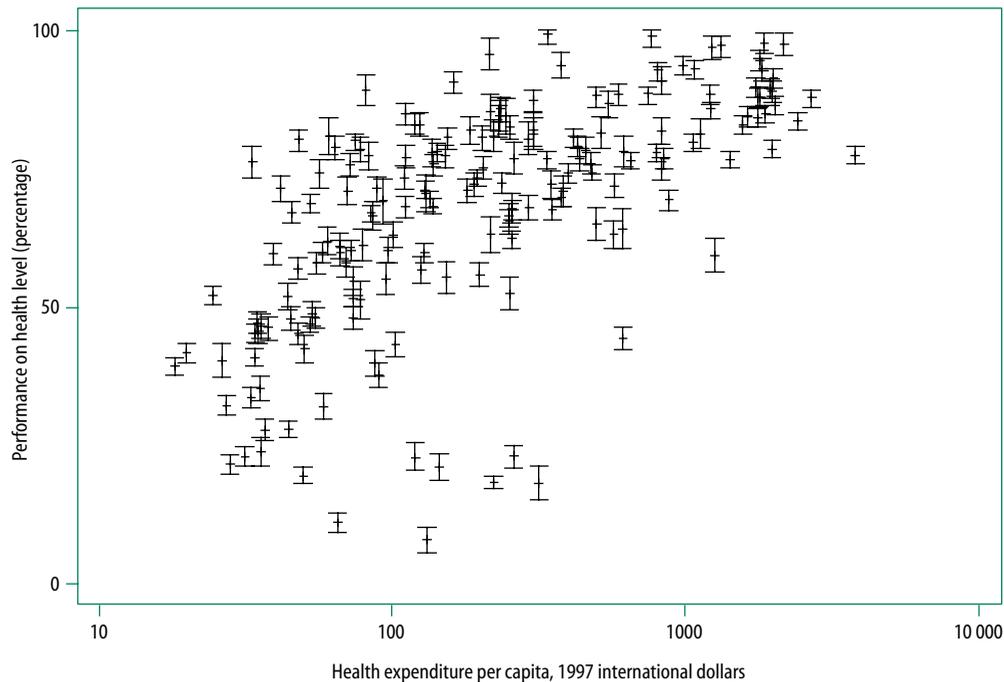
Annex Table 10 presents two indicators of health system performance. The first is based only on the average health status in disability-adjusted life expectancy (DALE) presented in Annex Table 5, comparing the frontier for that objective alone to a country's resource use and human capital. In this case, the upper and lower bounds between which performance lies are strictly comparable, and the measure can be compared to other estimates of what determines health outcomes. As with the measures of attainment, these values carry estimates of uncertainty. Figure 2.6 shows the estimated distribution of performance for all

countries with respect to DALE. Higher health expenditure is associated with better health outcomes, even when performance is judged relative to expenditure rather than absolutely. Very poor countries evidently suffer from other handicaps than low spending and low educational attainment. The few countries where spending is below \$10 per person per year seldom appear to achieve more than 75% of the life expectancy that should be possible, whereas most countries spending more than \$1000 achieve at least 75% of the possible. Higher spending is also associated with less variation in performance. Disturbingly large variations in life expectancy relative to spending and education occur at low and middle levels of expenditure where there is the greatest need to understand and reduce differences in achievement. A large part of the explanation is the HIV/AIDS epidemic: the 25 worst-off countries are all African nations suffering from a severe burden of AIDS. (Box 2.5 explains why the epidemic was not taken into account in defining the frontier of the possible.)

The second indicator in Annex Table 10 is based on the overall attainment measure presented in Annex Table 9 and assesses performance relative to the frontier defined for all five elements of achievement. The intervals around these values are much larger than for DALE alone because of the uncertainty surrounding the other components. These components also account for some considerable changes in the ranking, but the best performing systems still seem to be those of relatively rich countries and the worst off are predominantly poor and in Africa. Figure 2.7 presents the distribution of overall performance, which shows somewhat less variation than Figure 2.6: countries that perform poorly with respect to health alone sometimes compensate for this by doing better in responsiveness or financing or in dealing with health inequality. Nonetheless the rankings of the two performance measures are rather closely associated, with a small number of countries that do much better by one measure than by the other.

The belief that the system should be accountable for the level and distribution of attainment on the goals of health, responsiveness and fair financing, all relative to health

**Figure 2.6 Performance on level of health (disability-adjusted life expectancy) relative to health expenditure per capita, 191 Member States, 1999**



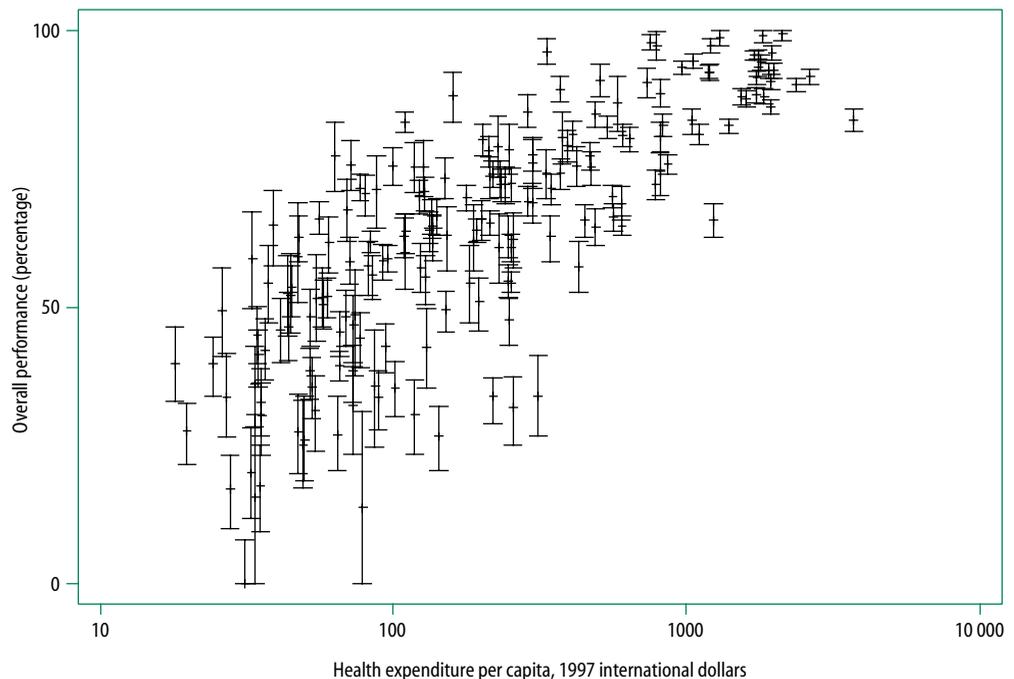
expenditure, will remain central in WHO's work to support health systems development over the coming years. From this issue, each year's *World health report* will contain more complete and better measures of countries' achievements, and WHO will support countries to strengthen local skills to analyse and improve health system attainment and performance.

## IMPROVING PERFORMANCE: FOUR KEY FUNCTIONS

Policy to improve performance requires information on the principal factors which explain it. Knowledge of the determinants of health system *performance*, as distinct from understanding of what determines health *status*, remains very limited. This report focuses on a few universal functions which health systems perform, as indicated in Figure 2.1 above, asking what it means for those functions to be discharged well or poorly and suggesting how they are associated with differences in achievement among countries. This helps to look at the health system overall, rather than building up from the component sub-systems, organizations or programmes, as is more common in evaluations of performance (18).

The service provision function is the most familiar, and in fact the entire health system is often identified with just service delivery. The classification here emphasizes that providing services is something the system *does*; it is not what the system *is*. Much of what is included in the financing function occurs outside what is usually considered to be the health system, as a process which happens to collect revenues and put them at the system's disposal. Treating fairness in financial contribution as one of the intrinsic goals of the system requires viewing the function partly as another of the tasks that the system *does*, rather than pas-

**Figure 2.7 Overall health system performance (all attainments) relative to health expenditure per capita, 191 Member States, 1997**



sively receiving money from somewhere else. It is the system which collects some of the funds directly, pools all that are pooled except for general taxation, and purchases goods and services. This means the system is at least partly accountable to society for how resources are raised and combined, and not only for how they are ultimately used.

Every health system makes some investments in creating resources, but these also are sometimes regarded as coming from outside the system itself. In the short run, the system can only use the resources created in the past, and often can do little to change even how they are employed. But in the long run, investment also is something the system does – and precisely because investments are long-lived, it has a responsibility to invest wisely. Relating achievements to total system expenditure may show that a system is performing badly precisely because what can be obtained from today's resources is needlessly limited by how resources were invested yesterday and the day before.

The fourth function is called *stewardship*, because the concept is well described by the dictionary definition: *the careful and responsible management of something entrusted to one's care* (19). People entrust both their bodies and their money to the health system, which has a responsibility to protect the former and use the latter wisely and well. The government is particularly called on to play the role of a steward, because it spends revenues that people are required to pay through taxes and social insurance, and because it makes many of the rules that are followed in private and voluntary transactions. It also owns facilities on trust from the citizens. Private insurers and practitioners, however, perform this function in only a slightly restricted degree, and part of the state's task as the overall steward or trustee of the system is to see to it that private organizations and actors also act carefully and responsibly. A large part of stewardship consists of regulation, whether undertaken by the government or by private bodies which regulate their members, often under general rules determined by government. But the concept embraces more than just regulation, and when properly conducted has a pervasive influence on all the workings of the system.

These functions are identifiable in widely differing health system structures (20, 1). At one extreme is a system in which functions are substantially combined in a single organization which raises, pools and allocates funds to a fairly monolithic group of service providers who are its own employees. The Norwegian health system resembles this type of structure, as did the British National Health Service prior to 1990. A system may instead have a high degree of "vertical" segmentation. Separate organizations such as the ministry of health, social security funds, the armed forces, charitable organizations, or private insurers may pay their own providers, raise and allocate funds and provide services, for non-overlapping populations. The health systems of much of Latin America bear some resemblance to this model, although patients often get care from two or more of the vertically separate organizations. A system could also have "horizontal" integration of each function – one organization performing it – but a different organization for each function. No system quite corresponds to this, because there is never a single bloc of providers, unless they are part of a fully integrated system. However, some systems such as that of Chile separate collection and pooling for a large share of resources, and employ a large number of providers under a single organization. At the opposite extreme from a monolithic organization is a system with separate institutions raising funds and paying providers under pluralistic provision arrangements in which few providers "belong" to the financing institution. The Colombian system, following the reforms introduced since 1993, looks somewhat like the latter.

Chapters 3, 4, 5 and 6 concentrate in turn on key characteristics of each of the four functions – service delivery, investment, financing and stewardship – and on some factors

affecting performance, examining patterns in countries at different income levels. The financing function obviously is most important for the goal of fairness in paying for the system, but how it is carried out also affects health outcomes and even has some effect on responsiveness. The service delivery function is most tied to health outcomes, but also matters greatly for responsiveness. And stewardship affects everything.

## REFERENCES

1. Murray CJL, Frenk J. *A WHO framework for health system performance assessment*. Geneva, World Health Organization, 1999 (GPE Discussion Paper No. 6).
2. Roemer MI. *National health systems of the world*. New York, Oxford University Press, 1991.
3. Centers for Disease Prevention and Control, Centro Latino-Americano de Investigación en Sistemas de Salud, and Pan American Health Organization. *Public health in the Americas*. Washington, DC, Pan American Health Organization, 1999 (unpublished document).
4. *World development report 1993 – Investing in health*. Washington, DC, The World Bank, 1993.
5. Acheson D et al. *Independent inquiry into inequalities in health*. London, The Stationery Office, 1998.
6. Gilson L, Alilio M, Heggenhougen K. Community satisfaction with primary health care services: an evaluation undertaken in the Morogoro region of Tanzania. *Social Science and Medicine*, 1994, **39**(6): 767–780.
7. Bassett MT, Bijlmakers L, Sanders DM. Professionalism, patient satisfaction and quality of health care: experience during Zimbabwe's structural adjustment programme. *Social Science and Medicine*, 1997, **45**(12): 1845–1852.
8. Hurst J. Challenges to health systems in OECD countries. *Bulletin of the World Health Organization*, 2000, **78**(6) (in press).
9. Donelan K et al. The cost of health system change: public discontent in five nations. *Health Affairs*, 1999, **18**(3): 206–216.
10. King S. Vaccination policies: individual rights v. community health. *British Medical Journal*, 1999, **319**: 1448–1449.
11. de Silva A. *A framework for measuring responsiveness*. Geneva, World Health Organization, Global Programme on Evidence for Health Policy, 1999 (unpublished paper).
12. Fabricant S, Kamara C, Mills A. Why the poor pay more: household curative expenditures in rural Sierra Leone. *International Journal of Health Planning and Management*, 1999, **14**: 179–199.
13. Nolan B, Turbat V. *Cost recovery in public health services in sub-Saharan Africa*. Washington, DC, Economic Development Institute of The World Bank, 1995.
14. Bennett S, Creese A, Monasch R. Health insurance schemes for people outside formal sector employment. Geneva, World Health Organization, 1998 (Current Concerns, ARA paper No. 16, document WHO/ARA/CC/98.1).
15. Friedman M. *Theory of the consumption function*. Princeton, NJ, Princeton University Press, 1957.
16. Bevan A. *In place of fear*. London, Heinemann, 1952.
17. Donabedian A, Wheeler JRC, Wyszewianski L. Quality, cost, and health: an integrative model. *Medical Care*, 1982, **20**: 975–992.
18. Cumper GE. *The evaluation of national health systems*. Oxford, Oxford University Press, 1991.
19. *Meriam Webster's deluxe dictionary*. Pleasantville, New York/Montreal, Reader's Digest, 1998.
20. Murray CJL, Kreuser J, Whang W. Cost-effectiveness analysis and policy choices: investing in health systems. *Bulletin of the World Health Organization*, 1994, **72**(4): 663–674.

## CHAPTER THREE

# *Health Services: Well Chosen, Well Organized?*

*Health services aim to protect or improve health. Whether they do so effectively depends on which services are provided and how they are organized. Resources should be used for interventions that are known to be effective, in accordance with national or local priorities. Because resources are limited, there will always be some form of rationing but prices should not be the chief way to determine who gets what care. Both hierarchical bureaucracies and fragmented, unregulated markets have serious flaws as ways to organize services: flexible integration of autonomous or semi-autonomous health care providers can mitigate the problems.*

## 3

## HEALTH SERVICES: WELL CHOSEN, WELL ORGANIZED?

### ORGANIZATIONAL FAILINGS

Just as the principal objective of a health system is to improve people's health, the chief function the system needs to perform is to deliver health services. The other functions matter partly because they contribute to service provision. It is therefore a major failing of the system when effective and affordable health interventions do not reach the populations that would benefit from them. Sometimes this happens because the providers have inadequate skills, or because of a lack of drugs and equipment: these are the consequence of failures of training and investment, as discussed in Chapter 4, or of purchasing, as discussed here and in Chapter 5. Sometimes services are not delivered to potential beneficiaries because of price barriers: this is the result of a failure to finance the services fairly, as discussed in Chapter 5. But often a failure of service delivery is due to dysfunctional organization of the health system, even when the needed inputs exist and financial support is adequate and fairly distributed. Such an organizational failing can result from the wrong arrangements among different parties involved in service delivery, which in turn creates perverse incentives and leads to mistaken choices about what services to provide, to whom to deliver them, or how to ration when it is not possible to meet everyone's needs or wants. This chapter considers how to choose which services to provide, how to organize provision and how to assure the right incentives for providers.

The complexities of organizing service provision are illustrated by the following example, which is not at all unusual. A poor young woman walks to a rural government health post with her sick baby. There is no doctor at the post, and there are no drugs. But a nurse gives the mother an oral rehydration kit and explains how to use it. She tells the mother to come back in a couple of days if the baby's diarrhoea continues. The nurse sees only half a dozen patients that day. Meanwhile, at the outpatient clinic of a community hospital about an hour's drive away, several hundred patients are waiting to be seen. Some are given cursory examinations by the doctors there and are able to obtain any prescribed drugs at the hospital dispensary. When the outpatient clinic closes, even though it is still early in the day, patients who have not been seen are asked to return the next day, without being given appointments. Some of the doctors then hurry off to work in a private "nursing home" or clinic to supplement their salaries.

The doctors' low pay and the absence of more qualified staff and drugs at the health post might be shrugged off as the consequences of spending too little. But a lack of resources cannot be blamed for the maldistribution of those resources between the health post and the hospital, the low productivity of the nurse, the under-utilization of the hospital when its

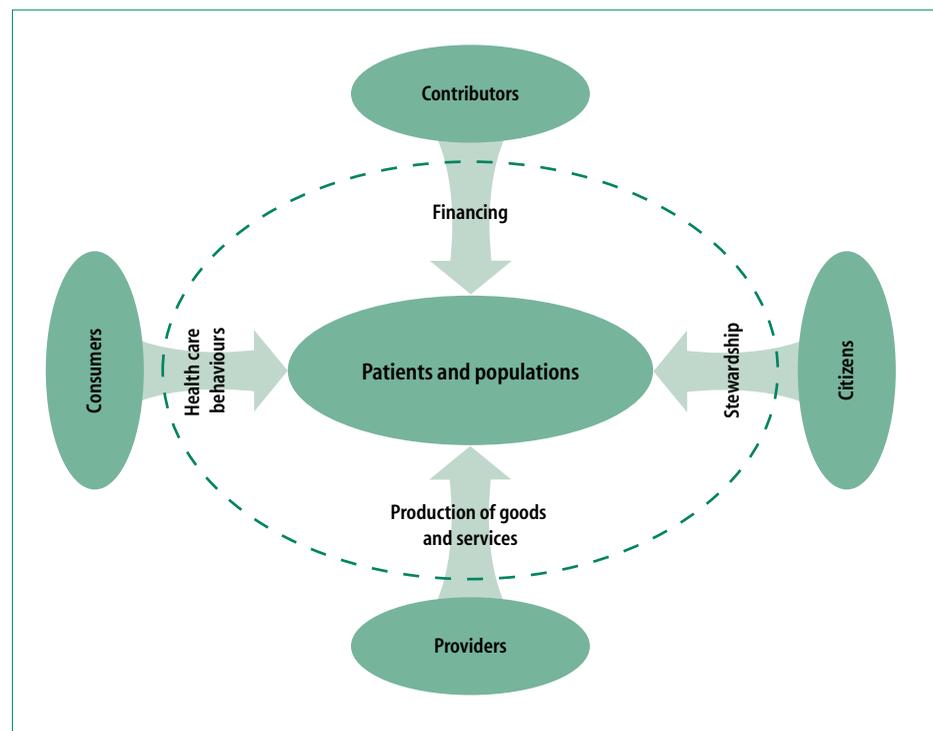
clinic closes early, the failure to have some doctors on duty over a longer interval, and the waste of people's time in waiting and then having to return another day because there is no system of appointments. These problems reflect failures of priority and of organization, both in initial investments and training and then in service delivery or the lack thereof. If the story has a happy ending for the mother and baby, it is only because the child was lucky to have diarrhoea and not malaria or some other condition the nurse could not recognize or could not treat, or requiring care which the mother would have to pay for out of pocket. Getting even limited care for free may also be the reason the mother goes to a public facility rather than to one of the private pharmacies or traditional healers, patronized by large numbers of people.

This chapter looks at how to set priorities for which services health systems should provide, and at the choices and mechanisms involved in rationing so as to make priorities effective. It then considers the organizational factors that help to make sure that the right services reach people at the right time.

## PEOPLE AT THE CENTRE OF HEALTH SERVICES

The story of the mother and baby illustrates another fact about health systems: service delivery is where people meet most directly, as providers and users of interventions. But people play more than those two roles, as Figure 3.1 indicates. At the centre of service delivery is the patient, in the case of clinical interventions, or the affected population, in the case of non-personal public health services. People are also consumers, because they behave in ways that influence their health, including their choices about seeking and utilizing health care. The consumer may be the patient, or someone such as a mother acting on his

**Figure 3.1** The multiple roles of people in health systems



or her behalf, or simply a person making choices about diet, lifestyle and other factors that affect health.

Sometimes the roles of consumer, patient and provider are all combined into one person and one moment, as happens when a woman gives birth with little or no assistance. Every minute, thousands of women across the world are giving birth. In countries where the attendance by trained staff is low (9% in Nepal, 8% in Bangladesh and Ethiopia, 5% in Equatorial Guinea, 4% in Gabon and Mauritania, 2% in Somalia), births usually take place in the presence of lay birth attendants or family members. Even when the delivery is by caesarian section with a trained provider, each woman must still actively participate in birth and the postpartum recovery.

Often the choices people make, particularly about seeking care, are influenced by the responsiveness of the system as described in Chapter 2. Utilization does not depend only on the consumer's perception of need or of the likelihood of benefiting from a service. Although marked differences exist between societies, the basic tenets of ethical provider-patient relations usually include similar elements of consent, confidentiality, discretion, veracity and fidelity (1). Calling the elements of dignity, autonomy and confidentiality that go into responsiveness "respect for persons" underscores the importance of people, and not simply patients, as the recipients of health services.

People also play the role of contributors to financing the system. Millions of poor people pay for all of the services they receive at the time they are ill. In health systems with fairer contribution arrangements, people who are not sick contribute most to financing the health system, through taxes or health insurance contributions, so that the contributor may or may not be the patient or the consumer. Finally, as citizens – and particularly as officials whose job it is to represent citizens and protect their interests – people participate in the system as stewards. In the same way that all four functions have to be carried out in order for the system to perform well, people have to play all these roles in order for the potential benefits to reach the patients and populations at the centre.

People act as providers, consumers, contributors and stewards of the health system during their adult working lives. In contrast, they can assume the role of patients at any time from before birth right up to death. The need to deliver services for people at all ages greatly complicates the choice of what services to emphasize and how to organize them, because people are exposed to different risks at different ages, and priority to any particular intervention is at least in part also a priority for a particular age group. These differences are what make a demographic transition – lower mortality and longer life – into an epidemiological transition – a change in the relative importance of different threats to health, particularly a shift from communicable to noncommunicable diseases.

Besides the variation with age, there are marked differences in disease patterns among regions, countries and specific population groups. For example, in Africa infectious diseases account for nearly 70% of the disease burden, as Annex Table 4 shows. In Europe, they account for less than 20%. The poor suffer more from infectious diseases than the rich (2), but over the next 20 years even the poor will be vulnerable to cardiovascular and cerebrovascular diseases linked to tobacco use (3). It may seem natural to focus health system choices on the causes that account for a large share of the disease burden, either because they affect large populations or because they cause substantial health loss for each victim.

However, all that health systems can actually *do* is to deliver specific services or interventions. Even if a first choice is made to concentrate on one or more particular diseases, it is still necessary to decide what to do – that is, which specific interventions to emphasize. The number of interventions available greatly exceeds the number of diseases, and the

appropriate strategy for disease control may depend on just one intervention or on a combination of several activities. To make matters more complicated, a given intervention may be effective against more than one disease or cause, because it works on a common risk factor or symptom. This is especially true of diagnostic activities: taking blood samples, or using X-rays or other imaging techniques may be appropriate for a great variety of problems. Thus, emphasizing an intervention, or investing in the inputs necessary for providing it, does not automatically focus effort on just one disease. Setting priorities also involves deciding what a particular intervention should be used for.

The range of diagnostic approaches and medical and surgical interventions for many conditions is extensive and likely to expand significantly over the coming decades. This means that services need to be designed and implemented so as to allow for innovation and adaptation to new health challenges and interventions, all the while responding to the needs of people who differ in age, income, habits and health risks. No health system can meet all those needs, even in rich countries. So either there must be conscious choices of what services should have priority, or the services actually delivered may bear little relation to any reasonable criterion of what is most important.

## CHOOSING INTERVENTIONS: GETTING THE MOST HEALTH FROM RESOURCES

The ancient Greeks believed that *Asclepios*, the god of medicine, had two daughters. One, *Hygieia*, was responsible for prevention, while the other, *Panacea*, was responsible for cure (4). While some preventive activities are applied to specific individuals – immunization is the clearest example – the distinction between prevention and cure or treatment corresponds closely to the difference between public health interventions directed to entire populations and clinical interventions directed to individuals. Since there is usually demand for the latter but there may not be any demand for the former, one of the principal tasks in choosing which services should have priority is that of balancing public health and clinical activities (5).

To require the health system to obtain the greatest possible level of health from the resources devoted to it, is to ask that it be as cost-effective as it can be. This is the basis for emphasizing those interventions that give the most value for money, and giving less priority to those that, much as they may help individuals, contribute little per dollar spent to the improvement of the population's health. It is the implicit basis of the measure of performance with respect to disability-adjusted life expectancy presented in Chapter 2 and Annex Table 10. So far as the level of health is concerned, the allocative efficiency of the health system could be enhanced by moving resources from cost-ineffective interventions to cost-effective ones (6). The potential gains from doing this are sometimes enormous, because the existing pattern of interventions includes some which cost a great deal and produce few additional years of life. For example, a set of 185 publicly-funded interventions in the United States cost about \$21.4 billion per year, for an estimated saving of 592 000 years of life (considering only premature deaths prevented). Re-allocating those funds to the most cost-effective interventions could save an additional 638 000 life years if all potential beneficiaries were reached. At the level of specific services, the cost per year of life saved can be as low as \$236 for screening and treating newborns with sickle-cell anaemia or as high as \$5.4 million for radionuclide emission control (7). In poor countries all the absolute numbers will be smaller, but the ratio between more and less cost-effective actions may still be very large.

Combining calculations of the cost with measures of the effectiveness of interventions and using them to determine priorities is a very recent development. Early work using such techniques in developing countries looked mainly at the cost-effectiveness of specific disease control programmes (8–13). This type of work expanded following publication of the *World development report* by the World Bank in 1993 (14) and subsequent work by WHO (15). Table 3.1 provides examples of interventions that, if implemented well, can substantially reduce the burden of disease, especially among the poor, and do so at a reasonable cost relative to results. Services can also be classified by their importance in the burden of disease of particular age and sex groups, and their cost-effectiveness for those groups (14).

Ideally, services with these virtues will also be inexpensive, so that they can be applied to large beneficiary populations and still imply reasonable total expenditures. However, there is no guarantee that low cost per life saved or healthy life year gained will mean low cost per person: some cost-effective interventions can be very expensive, with great variation

**Table 3.1 Interventions with a large potential impact on health outcomes**

Examples of interventions	Main contents of interventions
Treatment of tuberculosis	Directly observed treatment schedule (DOTS): administration of standardized short-course chemotherapy to all confirmed sputum smear positive cases of TB under supervision in the initial (2–3 months) phase
Maternal health and safe motherhood interventions	Family planning, prenatal and delivery care, clean and safe delivery by trained birth attendant, postpartum care, and essential obstetric care for high risk pregnancies and complications
Family planning	Information and education; availability and correct use of contraceptives
School health interventions	Health education and nutrition interventions, including anti-helminthic treatment, micronutrient supplementation and school meals
Integrated management of childhood illness	Case management of acute respiratory infections, diarrhoea, malaria, measles and malnutrition; immunization, feeding/breastfeeding counselling, micronutrient and iron supplementation, anti-helminthic treatment
HIV/AIDS prevention	Targeted information for sex workers, mass education awareness, counselling, screening, mass treatment for sexually transmitted diseases, safe blood supply
Treatment of sexually transmitted diseases	Case management using syndrome diagnosis and standard treatment algorithm
Immunization (EPI Plus)	BCG at birth; OPV at birth, 6, 10, 14 weeks; DPT at 6, 10, 14 weeks; HepB at birth, 6 and 9 months (optional); measles at 9 months; TT for women of child-bearing age
Malaria	Case management (early assessment and prompt treatment) and selected preventive measures (e.g. impregnated bed nets)
Tobacco control	Tobacco tax, information, nicotine replacement, legal action
Noncommunicable diseases and injuries	Selected early screening and secondary prevention

between one health service and another, for the same disease. This is clear in the case of malaria, where two interventions that are about equally cost-effective – chloroquine prophylaxis and two annual rounds of insecticide spraying – differ enormously in how much they would cost to apply to all the affected population of a low income African country (16). Cost differences are even greater for interventions against an infection.

The reverse is also true: health interventions can be cost-ineffective even when they do not cost very much and are intended to benefit large numbers of people. For example, many service providers continue to rely on antibiotics to treat viral illnesses, even though this is known to be ineffective. Even in rich countries, there is a need to ensure that the main output of health services remains focused on effective and affordable public health and clinical interventions. In low income countries, where the full range and cost of possible interventions significantly outstrip available resources, such wasteful practices deprive other patients of critical treatment.

Cost-effectiveness analysis, then, is essential for identifying the services that will produce the most health gain from available resources, but it has to be applied to individual interventions, not broadly against disease or causes. This requirement means that a large set of interventions needs to be evaluated. For all but the richest societies, the cost and time required for such an evaluation may be prohibitive. Moreover, such analysis, as currently practised, often fails to identify existing misallocation of resources because it focuses on the evaluation of new technologies and ignores the existing distribution of productive assets and activities (6).

Intervention costs can also vary greatly from one country, context, and intervention mode to another (17). A naive generalization could lead to serious mistakes in planning and implementing otherwise effective interventions. Even if they cover a relatively small number of interventions, studies in individual countries or populations are needed to avoid such errors. In Guinea, for example, 40 interventions have been studied. These were chosen partly on the basis of more general studies elsewhere, but with detailed local information to confirm what would really be most appropriate in that country (18).

Variations in cost and results among interventions are particularly relevant when a combination of several interventions may be suitable against a particular disease. To take the case of malaria again, at low levels of health expenditure in a country with a high burden of the disease, case management and prophylaxis for pregnant women would be very cost-effective and affordable (16). With more resources available, impregnated mosquito nets could be added – they would prevent more cases but cost more per unit of health benefit gained. A single estimate of cost-effectiveness of malaria control could lead to the wrong conclusion that malaria control is not affordable, for example if the estimate for a low income country is based on a programme combining all technically feasible options. In general, the most cost-effective combination of services depends on the resources available. That relation does not, of course, determine the appropriate level of expenditure on malaria control, which depends on what the country can afford, given its other health problems and priorities. In particular, there is no presumption that it should spend only the amount consistent with one or more of the cheapest interventions. Spending more and using a mixed strategy might yield much greater health gains.

Misuse of cost-effectiveness analysis could also lead to a serious underestimate of the actual cost of control if the estimate were based on the costs and effectiveness of a single type of intervention but multiple interventions were used. Many factors may alter the actual cost-effectiveness of a given intervention programme during implementation. These include: the availability, mix and quality of inputs (especially trained personnel, drugs, equip-

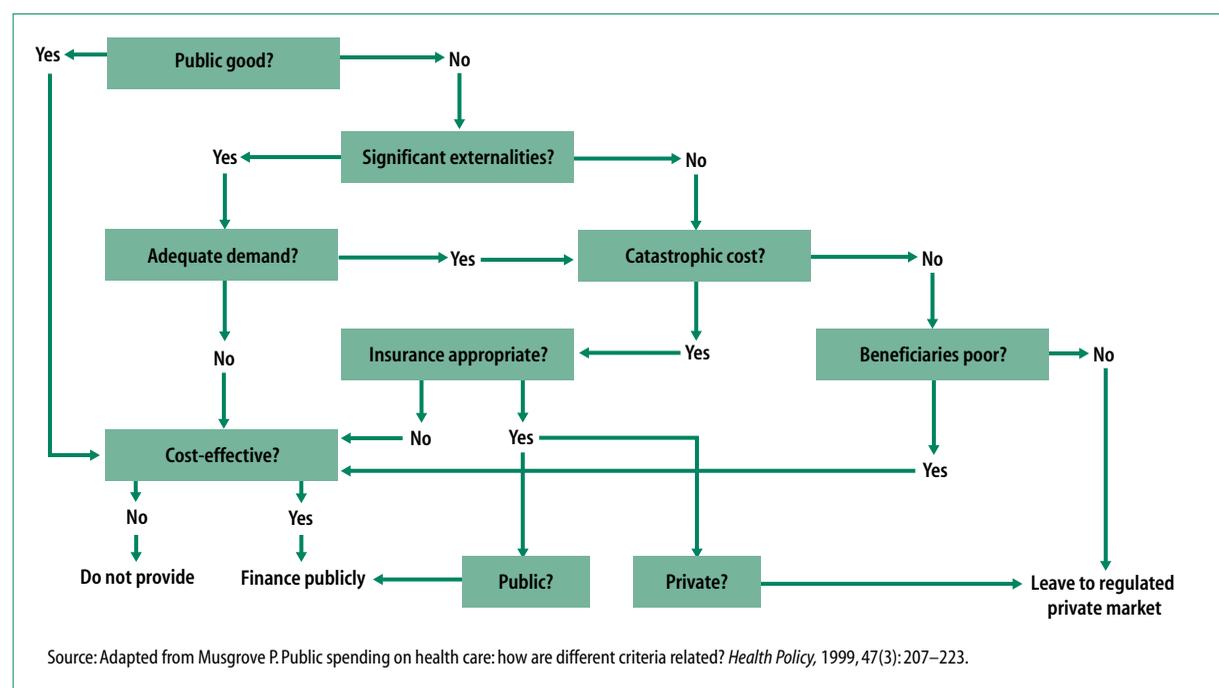
ment and consumables); local prices, especially labour costs; implementation capacity; underlying organizational structures and incentives; and the supporting institutional framework (17, 19).

All these obstacles imply that even on the sole criterion of cost-effectiveness, analysis of a health system's potential for getting more health from what it spends needs to begin with the current capacities, activities and outcomes, and consider what steps can be taken *from that starting point* to add, modify or eliminate services. This is likely to have profound implications for investment if little can be changed simply by re-directing the existing staff, facilities and equipment (20).

## CHOOSING INTERVENTIONS: WHAT ELSE MATTERS?

Cost-effectiveness by itself is relevant for achieving the best overall health, but not necessarily for the second health goal, that of reducing inequality. Populations with worse than average health may respond less well to an intervention, or cost more to reach or to treat, so that a concern for distribution implies a willingness to sacrifice some overall health gains for other criteria. More generally, cost-effectiveness is only one of at least nine criteria that a health system may be asked to respect. A health system ought to protect people from financial risk, to be consistent with the goal of fair financial contribution. This means that the cost matters, and not only its relation to health results, whether money is public or private. A health system should strive for both horizontal and vertical equity – treating alike all those who face the same health need, and treating preferentially those with the greatest needs – to be consistent with the goal of reducing health inequalities. And it should assure not only that the healthy subsidize the sick, as any prepayment arrangement will do in part, but also that the burden of financing is fairly shared by having the better-off subsidize the less well-off. This generally requires spending public funds in favour of the poor.

Figure 3.2 Questions to ask in deciding what interventions to finance and provide



Public money is also the principal, if not the only significant way to pay for public goods, interventions which private markets will not offer because buyers cannot appropriate all the benefits, and non-buyers cannot be excluded. The same is true for partly public goods with large externalities – that is, spillovers of benefits to non-users. Private demand for such services will generally be inadequate. Interventions of this sort are most important in communicable disease control, where treating one case may prevent many others, and especially where it is the environment, rather than identifiable individuals, that is treated. Analysts and decision-makers also correctly argue that resource allocation decisions affecting the entire health system must take into account social concerns, such as a priority for the seriously ill and for promoting the well-being of future generations. Figure 3.2 summarizes the choices for spending public or publicly mandated funds, showing how the different criteria should be considered sequentially and how they can be used to determine whether an intervention is worth buying or not. This way of setting priorities reinforces the emphasis on the two goals of health outcomes and financial fairness. It also emphasizes the importance of public health activities, by starting with interventions that are public or quasi-public goods.

Ignoring these other criteria and using only disease burden and cost-effectiveness as a method for determining priorities can lead to a “race for the bottom of the barrel” among advocates of different interventions, each trying to prove that their programme achieves a greater benefit or costs less than other programmes, sometimes without considering the full range of complicating factors. This often leads to underestimates of the real cost of programmes and their subsequent failure during implementation because of resource shortages.

Too narrow an approach also ignores the important role that the public sector should be playing in protecting the poor and addressing insurance market failure – the tendency of insurance to exclude precisely those people who need it most, because they are at greater than usual risk of ill-health. Many families will be faced at some time with a health problem of low frequency for which there is an effective but high cost intervention. Those who can afford it will turn to the private sector for the needed care. But without some form of organized insurance this option is usually too expensive for the poor who will turn to public hospitals as a place of last recourse. Often this leads to inappropriate and excessive use of hospital care, and it undermines the financing function that health systems should be playing.

Actual health systems always deliver services that correspond to a variety of criteria. The frontier of the possible which defines relative performance reflects this fact, since it is based on actual outcomes relative to health expenditure and human capital. A health system designed and operated solely to pursue cost-effectiveness might be able to achieve much longer average life expectancy or more equality or both, but it would correspond much less to what people want and expect.

What makes it particularly difficult to set priorities among interventions and beneficiaries of health services is that the different criteria are not always compatible. In particular, efficiency and equity can easily be in conflict, because the costs of treating a given health problem differ among individuals, or because the severity of a disease bears little relation to the effectiveness of interventions against it or to their cost. Cost-effectiveness is never the *only* justification for spending public resources, but it *is* the test that must be met most often in deciding which interventions to buy. And it can be set aside only when costs are low and the beneficiaries are not poor, so that they can make their own judgements about the value of a particular purchase and the market can be left to supply it; or when protection from

catastrophic cost is the overriding consideration and prepayment can protect against that risk. Determining the priorities for a health system is an exercise that draws on a variety of technical, ethical and political criteria and is always subject to modification as a result of experience in implementation, the reaction of the public, and the inertia of financing and investment (21).

## CHOOSING INTERVENTIONS: WHAT MUST BE KNOWN?

Setting priorities realistically requires a great deal of information, starting with epidemiological data. Major progress has been made recently in understanding global health and disease patterns (14, 15, 22), including analysis of risk factors which influence several diseases at once. The most significant of such risk factors are malnutrition in children, and poor water and sanitation practices. Other major risk factors include unsafe sex, alcohol, indoor pollution, tobacco, occupational hazards, hypertension and physical inactivity. The public health services in a given country should attempt to deal with such preventable risk factors, taking account of local contexts. For example, the origins of malnutrition vary greatly from one country to another and from one region to another. In sub-Saharan Africa and south Asia, the problem is often a combination of micronutrient deficiency and absolute shortage of calories. In central and eastern Europe, malnutrition is often “poor calories” rather than a “lack of calories” – a diet too high in fat and refined starch. Public health activities will therefore vary, depending on local risk factors and diseases conditions.

Although there are good data on national patterns of risk and disease today, few countries break this information down sub-nationally by income level, sex or vulnerable groups, such as the handicapped, minority ethnic populations, and the frail elderly. Even fewer countries have information on the health-seeking behaviour of those groups or their utilization of health care facilities. Without such information, the effectiveness of interventions is difficult to assess, as the same intervention may have very different effects when applied to different populations.

Governments need to know how to influence the health-seeking behaviour of target groups in need of care. For example, intergroup variations in under-5 mortality are particularly large in Brazil, Nicaragua, and the Philippines, whereas in Ghana, Pakistan, and Viet Nam these differences are much smaller. This shows the need for a greater emphasis on equity in providing health services in the former countries (23). And there are often significant differences in the utilization of preventive and clinical medical attention from one intervention to another, in the same country. In Peru, differences between the rich and poor are far greater with respect to attended deliveries than with respect to immunization (24), largely because of the much higher cost of deliveries.

A key recommendation for policy-makers is to collect and combine data on risk factors, health conditions and interventions with data from household and facilities surveys, focus groups and other qualitative methods, and academic studies, since global and national aggregate data may not reflect local needs. Public health and clinical services should be customized to respond to the latter, and should allow for innovative adaptation during implementation. While gathering and analysing such data is more difficult in the very poorest countries which need this type of analysis the most, the methods are becoming routine and more easily used even at low incomes (25).

The following steps will make health systems more likely to produce effective interventions at an affordable cost, especially for needy populations.

- First, there should be an ongoing detailed assessment of underlying risk factors, disease burden, and utilization patterns of the target populations.
- Second, global information on the cost and effectiveness of interventions, as well as intervention strategies and practice patterns, should be adapted to local prices and local contexts.
- Third, all countries need explicit policies to ration interventions and to ensure that limited resources are spent in identified high priority areas. How to achieve this is taken up next. Few countries have clinical protocols that can be used to standardize practice patterns and match known priority interventions with needs. Fewer still have the means to enforce such guidelines in privately financed provision.
- Finally, none of these steps will matter unless the quality of service delivery is assured.

## ENFORCING PRIORITIES BY RATIONING CARE

Stating priorities is one thing; actually delivering the supposedly most valuable services at the expense of other services is another thing. Markets solve this problem through rationing by price, which means that who gets what goods and services depends not only on how much those goods and services are valued by people, but on who has the means to buy them. Priorities are not set by anyone but emerge from the play of the market. As indicated, this is almost the worst possible way to determine who gets which health services. Every health system therefore confronts the question of what other means to use, when resources are inadequate to needs or wants.

In low-income countries, the difficulties involved in setting priorities and rationing services are extreme. The HIV/AIDS epidemic kills over two million people in Africa every year – more than 10 times the number that perish in wars and armed conflict during the same period. The health services of many low income countries in south Asia and sub-Saharan Africa have been burdened in recent years by this epidemic. In this case, health systems are faced with a long-term problem. Difficult choices have to be made on how resources should be allocated to cover AIDS prevention campaigns, and care for people with AIDS, while maintaining other essential health services. This problem is chronic, and quite different from the need to ration non-urgent care when the system is temporarily burdened by a short-lived epidemic of disease or the results of a natural or man-made disaster. Then emergency services get priority, elective procedures are delayed, and the system concentrates on the epidemic until it is sufficiently under control that business as usual can be resumed.

The most common chronic approach to rationing care is to impose strict expenditure controls that do not try to target any specific disease group or broad category of interventions but simply limit budgetary obligations to affordable levels. This technique has been most commonly used in health systems with global budget financing and leaves it to the budget-holder to ration care. It has been used in the pre-1990 British National Health Service and the ministries of health of many low income countries. Other cost-containment techniques are now being used with varying degrees of success in many European Union countries and some developing countries (26).

The major disadvantage of this approach is that, in low income countries, it usually leads to a degradation of overall standards and quality of care. If resources are in the hands of the better-off, there may be a failure to target vulnerable groups. The available budget is usually captured by the politically strongest providers, such as specialists and hospitals,

rather than being used according to the needs of the population. Thus, in many low income countries, an approach based solely on expenditure control leads to the exclusion of large segments of the population from access to organized care.

A second approach is to ration explicitly, following priorities which were set according to some predetermined criteria, as discussed above. This approach, first introduced in the mid-1980s, has now been partially implemented in the Netherlands, New Zealand, Norway, Sweden and Oregon (USA) (27). All use a combination of social, political and cost-effectiveness criteria. Since 1993, several developing countries have tried to introduce intervention packages, variously described as including “essential” or “basic” or “core” interventions that are affordable within each country context (28). Mexico was the first country to design and adopt such a package (29). Bangladesh, Colombia and Zambia have also begun implementation.

The explicit priorities established through this process are a major improvement over the traditional passive cost-containment approach. One serious disadvantage is that in real life, providers are faced with demand for services that are not included in the defined benefit package. They usually react to this demand in one of two ways – by cross-subsidizing the excluded activities through the budget received to pay for the defined benefit package; or by charging extra for the additional services. The first leads to a financing shortfall for the defined benefit package. The second leads to increases in out-of-pocket expenditure and erosion in financial protection. Attempts to curtail such behaviour by providers have been largely unsuccessful.

Another problem is that there are “limits to rationality” (30), particularly if rationality is identified purely with cost-effectiveness. Politicians, providers and the public care about all the criteria discussed above, and may be very sceptical of the estimates underlying allocative choices. The success of explicit priority setting depends on the acceptance and support of providers and consumers.

Even within the set of services financed by prepayment, and particularly those financed by public or publicly mandated funds, there is no clearly best way to ration care. Figure 3.3 illustrates four simplified approaches, based on a combination of what services cost per individual treated or affected, and how frequently the service is likely to be needed. In general, very costly services are seldom needed, while there is much more frequent need for a variety of interventions with intermediate costs. The upper curve in each panel of the figure shows what the demand for different services might look like in the absence of any form of rationing – that is, if every need were expressed as a demand and there were no price or other barriers to obtaining care. That represents the most that the health system might want or try to deliver.

One way to limit what is actually delivered is to exclude all or most of the rare but very expensive services – to cut off the right-hand tail of the distribution of needs. This is relatively common in private insurance, either by explicit exclusion of services or by risk selection of potential clients so as to reduce the likelihood of those services. This may be, but need not be, consistent with cost-effectiveness, and it is almost a necessary form of rationing in systems with very limited resources. But it maximizes people’s exposure to financial risk if the intervention can be had by paying out of pocket, or to catastrophic health losses if the service is simply not available at all.

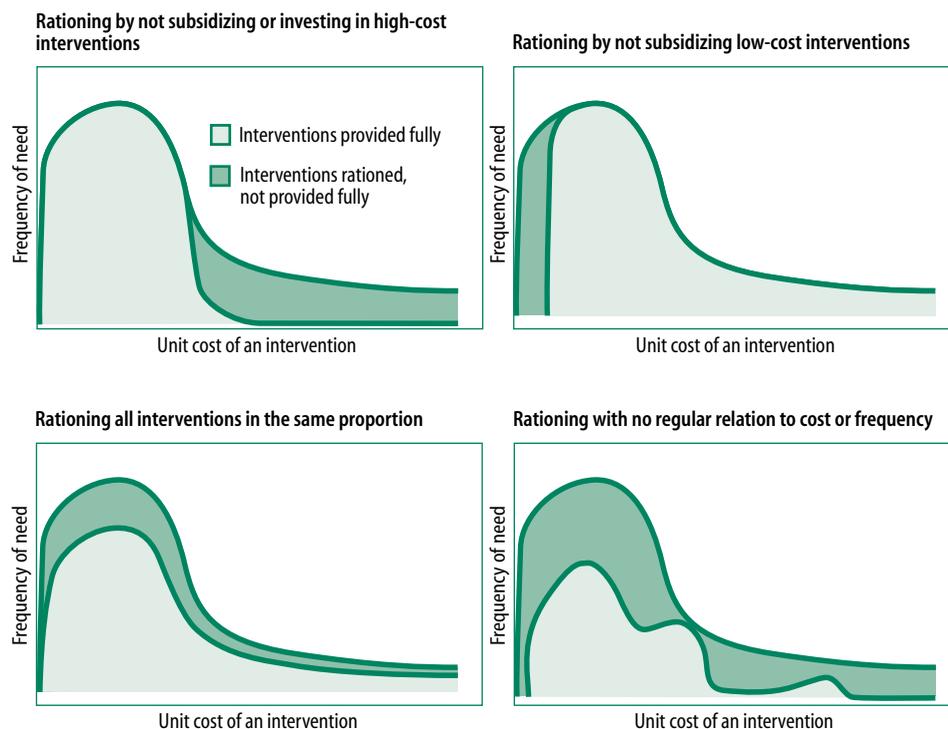
The opposite approach is to exclude common but very inexpensive services from prepayment schemes and in effect require that they be paid out of pocket – that is, to cut off the left-hand end of the distribution. This is likely to save administrative costs, but may or may not represent substantial overall cost saving. As a general rule, prices should not be the

main instrument of rationing, and low prices paid by the non-poor present a relatively minor problem. The difficulty with this approach is clear: it exposes the poor to risks that would be acceptable for the non-poor, and so worsens inequality in financial contribution. Rationing may need to be differently conducted for the poor than for the rest of the population, if prices are to play any role.

A health system could also try to ration all services in the same proportion, giving everyone who needs it the same likelihood of obtaining care independently of its cost or of how many other people need the same intervention. There is little to be said for this way of delivering less care than people need, since it does not respect any of the criteria discussed above. At best it represents an attempt to spread the frustration of not obtaining care more or less equally, but that does not even correspond to equality of responsiveness. It may be the response of a system under pressure and with no clear guidance as to the relative importance of different services.

The last panel of Figure 3.3 corresponds to explicit priority setting, so that rationing is much more severe for some services than for others. Only if this happens are nominal priorities really being enforced so as to affect service delivery. And only if the priorities are chosen according to some appropriate criteria can rationing, however it is enforced, actually contribute to better health system performance.

**Figure 3.3 Different ways of rationing health interventions according to cost and frequency of need**



## AFTER CHOOSING PRIORITIES: SERVICE ORGANIZATION AND PROVIDER INCENTIVES

Priority setting is generally considered a public sector exercise, particularly concerning the proper use of public or publicly mandated expenditure. It does not matter for this purpose whether the delivery of services is public or private, nor how providers are paid. What matters is that by contracting with private providers or reimbursing them through public insurance, the government can assure that its priorities are respected even where it does not provide the services. In the sphere of private, voluntary financing of services there generally are no explicit priorities: that part of the health system responds to demands rather than to needs. It is important to take into consideration the impact of out-of-pocket spending on the poor due to increased demand in the private sector for interventions that are not in the public package. But some priorities can nonetheless be enforced through regulation, as for example by requiring all private insurance policies to include a package of essential services or by limiting the degree to which private providers or insurers can select patients or clients on the basis of risk. These are among the tasks of stewardship discussed in Chapter 6. More generally, since it is ultimately providers who do or do not deliver the priority services, rationing requires “careful governance of the agents” who act for patients and assess their competing health needs (31).

Given a list of priorities, and given one or more mechanisms for rationing care, the way services are actually delivered – who benefits from which interventions, how efficiently they are provided, how responsive the system is – can still differ markedly from one health system to another. These differences reflect the fact that while providers may be urged or enjoined to deliver particular services, and public budgets and regulations are designed to reinforce those choices, there is still a variable latitude for providers themselves to decide whom to treat, for what, and how. Just how much latitude providers should have is one of the crucial questions for a health system. The outcome depends on organizational and institutional characteristics, which together determine some of the fundamental incentives to which providers respond.

The relationship between organizations, institutions, and interventions parallels that between the players, the rules and the objects of a game. *Organizations* are the players – for example, individual providers, hospitals, clinics, pharmacies, and public health programmes. *Institutions* are the rules (formal rules and informal customs) – the socially shared constraints that shape human interactions, along with the mechanisms by which these rules are enforced. The key institutions that affect the service delivery system include rules relating to stewardship (governance, information dissemination, coordination, and regulation) and purchasing. *Interventions*, in the sense of services or activities as described above, are the objects of the game and include clinical treatment, public health measures, and health-promoting intersectoral actions (32). *Incentives* are all the rewards and punishments that providers face as a consequence of the organizations in which they work, the institutions under which they operate and the specific interventions they provide.

Both among and within countries there are marked differences in all these features, reflecting the complexity of the production process for health interventions and the variations in culture and tradition. The characteristics that exert the most powerful influence on clinical and public health services are the organizational structures or forms, the service delivery configurations, the organizational incentive regimes, and the linkages among services. As emphasized in Chapter 1, health services deal with an asset – the human body – that is very different from those that other economic activities deal with. Nonetheless there

are some aspects of how health services are produced that do not differ greatly from the production of other services. Evidence of the importance of these factors is slowly growing as a result of progress made in applying systems analysis and organizational theory to health services (33–35).

## ORGANIZATIONAL FORMS

Health services can be organized in three fundamentally different ways – via hierarchical bureaucracies, through long-term contractual arrangements under some degree of non-market control, and as direct, short-term market-based interactions between patients and providers (36). These arrangements are independent of whether ownership is public or private. For example, the ownership of services that are organized as hierarchies can be public, as in the extensive network of public health, hospital, and ambulatory clinics that are part of the Turkish Ministry of Health service delivery system and that of many other countries. But they can also be private, as in a United States health management organization like Kaiser Permanente. Such private entities often suffer from many of the same bureaucratic rigidities as public ones. Likewise although market-based interaction between providers and patients is most common in the private sector, short-term market exchanges in the form of user fees are pervasive in the public sector in many low income countries.

India provides examples of all three organizational forms. The services delivered by government are hierarchical, with providers who are employed directly. At the other extreme are direct, market-based, non-contractual interactions between the population and providers. These include both private providers *per se* and informal fee charging in public facilities: 80% of total health care spending takes place in this domain. In between are several forms of contractual arrangement. One type comprises long-term contracts between the public sector and nongovernmental providers (both non-profit and for-profit). This arrangement is used predominantly for treating patients suffering from cataract and, by increasing the number of providers that are financed publicly, has allowed for a large expansion of surgery to prevent blindness, particularly among the poor. Another contractual arrangement characterizes private insurance, which may or may not be publicly regulated. The client has one kind of relation with an insurer, which in turn has a different relation – one that may or may not be contractual – with providers.

Each of these ways to organize health services has its strengths and weaknesses in various contexts and when applied to different types of population-based and clinical services. When a strongly coordinated approach is needed, as was the case for example during the postwar (late 1990s) reconstruction of the health service in Bosnia and Herzegovina or during an outbreak of cholera, hierarchical controls are better. Largely inspired by experiences such as the British National Health Service and the difficulty of addressing health problems through markets alone, many low and middle income countries have, over the past 50 years, established state-funded health care systems with services produced by a vertically integrated public bureaucracy. This has led to improved access to health care for millions of people and underpinned many successful public health programmes.

But hierarchical bureaucracies also have some serious shortcomings when it comes to the provision of health services. These shortcomings have become more apparent in recent years (37, 38). Bureaucracies are vulnerable to capture by the vested interests of the bureaucrats and providers who work in them. They are often not as effective in downsizing or reorienting priorities as they are in expanding capacity and adding services. And they are often associated with many of the same shortcomings as private markets in terms of abuse

of monopoly power (such as the collection of rents in the form of informal charges) and information asymmetry. Over time, many of the hierarchical service delivery systems have become excessively rigid, with inefficient processes producing low-quality care that is unresponsive to the needs and expectations of the populations and individuals that they serve. This has been the motive for many recent reform efforts, as described in Chapter 1.

Where there is a call for innovation and flexibility to respond to specific needs, as in the development of new drugs and equipment, markets are better. But direct market interactions between patients and providers in the health sector have the major disadvantage of exposing individuals to the financial risks of illness unless the financial resources are adequately pooled. And it is difficult or impossible to assure that such transactions respect any priorities among interventions and patients that the health system is trying to implement.

Because of the disadvantages of both rigid hierarchies and out-of-pocket payment in the health sector, countries throughout the world are today experimenting with long-term contracts to achieve the combined advantages of greater flexibility and scope for innovation while maintaining overall control over strategic objectives and financial protection. There is already some analysis of experiments with contracting for service provision in low and middle income countries (39), and much effort has also gone into drawing lessons from the better documented instances, particularly in the United Kingdom, which may also be relevant elsewhere (40).

## SERVICE DELIVERY CONFIGURATIONS

Health services, like many other forms of production, can be implemented in more dispersed or more concentrated configurations, or in hybrid arrangements that combine some concentrated with some dispersed elements (41). Dispersed service configurations are usual for activities which do not benefit from economies of scale – unit costs are no lower for large than for small production units – such as primary care, including the integrated management of childhood illness; pharmacies; dental offices; field-based implementation of public health programmes; counselling; social work; and community and home-based care. Such ambulatory services usually involve a fairly broad range of activities of varying degrees of complexity, such as the management of common clinical and nonclinical activities by individuals or small teams of people.

Dispersed, competitive production by small producing units works well wherever markets are a satisfactory way to organize output. It is less successful in health, for all the reasons that markets work more poorly for health care. However, attempts to offset market failings by integrating such dispersed activities into a hierarchical bureaucratic structure have almost always run into problems of staff motivation and accountability. Close supervision is difficult to implement, while excessive control is detrimental. A more successful approach has been to establish a contractual relationship that relies on professional reputation, and a strong sense of commitment and responsibility. Such contractual relationships have a long history of success in countries such as Denmark and Norway, and have recently been tried successfully in Croatia, the Czech Republic, and Hungary.

Concentrated service configurations are common for activities such as hospital care, central public health laboratories, and health education facilities, which do benefit from economies of scale – lower costs with larger size – and scope – lower costs from undertaking a variety of activities (42, 43). These interventions are highly specialized and expensive, and require large teams of people with a wide range of skills. Some require continuous observation (for surgical treatment and care), and highly controlled sterile conditions (for

surgical and burns units). Accountability can usually be enforced through direct observation of outputs or outcome. Most personnel can be employed as regular or part-time staff, rather than under the contractual relationships that appear to be better for dispersed activities. Countries have been more successful in integrating these services into hierarchical public bureaucracies but pay the price of the disadvantages of this organizational form.

There is both an upper and a lower efficiency boundary for concentrated service configurations. At the upper end, the large 1000 to 2000 bed hospitals and huge public health laboratories in central and eastern Europe were characterized by over-specialization, low productivity and low quality of care (44). At the lower end, there are also considerable efficiency and quality problems when facilities that perform specialized care are too small. Cottage or district-level hospitals with 20 to 50 beds are common in many low and middle income countries, such as Ethiopia, Morocco, and Turkey, especially in rural regions and in the private sector (45). Often they have low bed-occupancy rates and the staff do not see a sufficient volume of patients to maintain the clinical skills needed to treat rarer conditions. They may deal well with more common conditions, but then they must be integrated into a referral system that can treat more difficult or unusual ailments.

Hybrid service configurations fall somewhere between these two extremes. Many of the activities with a large potential impact on outcomes (shown in Table 3.1) are implemented in this form. Programmes to control infectious diseases such as malaria, tuberculosis and HIV/AIDS benefit from the planned coordination of some of their strategic elements at the national level. Yet their implementation can sometimes be more effective when carried out under contractual relationships with local providers than when implemented as vertical programmes isolated from other ambulatory services. For example, the implementation of the integrated management of childhood illness in Egypt requires close national coordination of activities such as immunization, malaria control and iron supplementation, but implementation would be impossible without local providers with the broad range of skills needed, for example, to treat acute respiratory infections, diarrhoea, and childhood illnesses.

This latter example highlights a key challenge in health service delivery. That is, to balance the need for broad policy oversight with sufficient flexibility so that managers and providers can innovate and adapt policies to local needs and contexts in a dynamic way. Population-based and clinical health services that can be refashioned through negotiation and adapted during implementation at the discretion of agencies and their staff are more responsive to the health needs and non-health expectations of the population than those that are implemented through rigid centralized bureaucracies (46–49). This is consistent with the relations between responsiveness and service characteristics described in Chapter 2. But this approach may lead to outcomes quite different from those intended at the outset. The more focused managers and staff are in pursuing a clear mandate, the more likely it is that broader policy objectives will be achieved without having to resort to rigid hierarchical structures for control (50).

## ALIGNING INCENTIVES

Service providers need flexibility, not for arbitrary purposes, but so that they can respond to well-defined incentives – that is, so the incentives defined by organizational and institutional arrangements can be effective instead of being frustrated by rigidities. The growing awareness of the structural nature of problems in hierarchical service delivery systems has led policy-makers in many countries to examine the incentive environment of organizations and alter the distribution of decision-making control, revenue rights, and financial

risk among the different participants, as analysed in *The world health report 1999* (16).

There is a wide range of ways to change the organizational incentive regime of health services. In many Latin American countries, including Argentina and Brazil, decentralization has led to a shift in decision-making control and often revenue rights and responsibilities from central to lower levels of government. The devolution of central control to provinces in Sri Lanka is another form of decentralization. The creation of semi-autonomous hospitals in Indonesia shifted decision-making and control even further down the line to the level of facilities. In Hungary, during the early 1990s, general practitioners were transformed from civil servants into semi-autonomous practitioners on contract with local governments and the newly created National Health Insurance Fund.

In each of these examples, there is a change in one or more organizational incentives that exerts a powerful influence on how the organizational unit in question behaves, be it a province, region, district, or individual provider unit such as a hospital or ambulatory clinic (51). Figure 3.4 shows the relation between the organizational forms discussed earlier and the following five incentives.

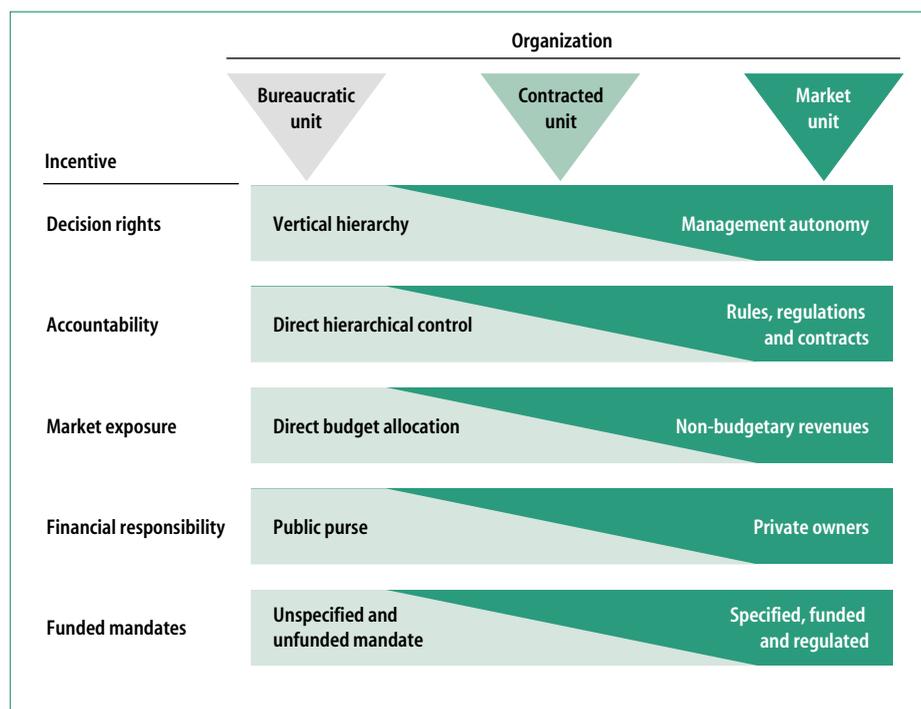
- The degree of *autonomy* (decision rights) that the organization has *vis-à-vis* its owners, policy-based purchasers such as insurance funds, the government, and consumers. Critical decision rights include control over input mix and level, outputs and scope of activities, financial management, clinical and nonclinical administration, strategic management and market strategy (where appropriate).
- The degree of *accountability*. As decision rights are delegated to the organization, the ability of governments to assert direct accountability (through the hierarchy) is diminished. When autonomy increases, accountability must be secured by shifting from hierarchical supervision to reliance on monitoring, regulations, and the economic incentives embedded in contracts.
- The degree of *market exposure* or revenues that are earned in a competitive way rather than through a direct budget allocation. Market participation need not imply out-of-pocket financing; it is preferable for provider organizations to compete for prepaid revenues. When governments bail out organizations that run deficits or are indebted as a result of weak technical performance, they undermine the impact of market exposure.
- The degree of *financial responsibility* for losses and rights to profit (retained earnings and the proceeds from the sale of capital). This determines the financial incentive for managers and staff to economize. Under increased autonomy they, rather than the public purse, become the “residual claimant” on revenue flows, but such claims must be clearly spelled out and regulated.
- The degree of *unfunded mandates*. Where the share of total revenues earned through markets is significant, organizations are at financial risk because of the unrecoverable costs associated with requirements for which no funds are provided, such as care for the poor or very sick. Organizational reforms that increase autonomy should therefore be accompanied by complementary reforms in health financing to protect the poor. Chapter 5 discusses some recent examples in Latin America.

How far countries can safely go in pushing service provision away from hierarchical control and towards an incentive environment (the right of the spectrum in Figure 3.4) depends on the nature of the services and the capacity to create accountability for public objectives through indirect mechanisms such as regulation and contracting. There is no single blueprint for a successful service delivery system. But countries such as Canada (52)

that have succeeded in creating a more coherent framework for these three organizational characteristics perform better than countries such as the United States (53) where there are many conflicting signals because market incentives are very strong in some places and more tightly controlled in others.

The coherence of organizational incentives is especially important in the hospital sector because of the central role of these organizations in service provision. Countries that have introduced consistent objectives and that have aligned the five organizational incentives appear to have been more successful than countries that have ended up with conflicting objectives and incentives regimes. For example, in Singapore, the public hospitals have been given considerable autonomy over management decisions ranging from procurement to personnel (54). Accountability is now enforced through contracts rather than hierarchical controls. The hospitals compete with each other for patients and can keep any surpluses they generate through savings. And there is an explicit subsidy scheme for low income groups, although cross-subsidies are still needed to cover some unfunded mandates. Follow-up assessments indicate that the reforms have succeeded in improving responsiveness to patients and efficiency in resource management, while protecting poor patients against opportunistic behaviour by hospitals trying to increase their revenues. In Indonesia, the degree of autonomy is much less but the various incentives are nevertheless more balanced than in New Zealand and the United Kingdom where there has been less policy coherence across the five organizational incentives (39, 55, 56). Hospitals are without question the most complex organizations involved in service delivery, and their role has been undergoing rapid change as new procedures shift the balance between inpatient and ambulatory care and as financial pressures have increased (57). How to organize hospital services and how to integrate them with other providers is perhaps the hardest question a service delivery system faces.

**Figure 3.4 Different internal incentives in three organizational structures**



One way that many countries have tried to increase market exposure of hospitals is to “outsource” or “unbundle” some hospital activities. Experience so far in this area has been mixed. For example, there has been some success in outsourcing the maintenance of medical equipment in Thailand, management services in South Africa, and routine custodial, dietary, and laundry services in Bombay. Most of these activities benefit from the efficiency gains that can be provided by external suppliers that specialize in a given service. But with few exceptions, outsourcing is much more difficult with clinical services because of loss of strategic control over part of the production process, cost shifting, and difficulties in monitoring the quality of the outputs (58).

Many public health interventions, such as malaria control programmes, nutrition programmes in Senegal, and reproductive health programmes in Bangladesh, are now carried out through long-term contracts with nongovernmental providers rather than rigid vertical programmes under a central hierarchical bureaucracy. And there has been a marked increase in the autonomy and privatization of general practitioners, dentists, pharmacists and other ambulatory health care workers in central and eastern Europe, with both good and bad consequences.

As in the case of hospitals, ambulatory services that are made autonomous perform better when there are minimal conflicts between the objectives and organizational incentive regimes. Table 3.2 provides some examples of organizational incentives for ambulatory

**Table 3.2 Examples of organizational incentives for ambulatory care**

Organization affected	Country examples
<p><b>Local or district teams that manage several clinical facilities and public health services</b></p> <p>Includes district level ministry of health offices and municipal councils. Changes in organizational incentives are often modest and mostly related to decision rights over budget and staff. Financial risk remains limited. Actual degree of market exposure may be greater than intended when user fees are significant.</p>	<p><b>Finland:</b> municipalities own and manage health centres, employ staff, raise taxes and set fees.</p> <p><b>Philippines:</b> decentralization of responsibility for primary health care (and other social services) to local governments in 1993. Assets, staff and budgets transferred to local level. Ministry of Health (MoH) set up community health care associations along with each local government unit. Health workers now report to local government, not to MoH. Supervision by MoH has become more difficult.</p> <p><b>Zambia:</b> the Central Board of Health (CBoH), the executive arm of the MoH, now contracts through annual district plans with independent district health boards/ district teams. Districts have gained greater control over their non-salary recurrent budget. But staff are mostly still employed by the civil service. This is changing as new graduates are hired by districts and unskilled staff are recruited locally. Accountability to CBoH is retained through sanctions if agreed performance targets are not met. Income from user fees is retained by facilities.</p>
<p><b>Individual facilities</b></p>	<p><b>Belarus:</b> polyclinics now receive their own budget and can retain a proportion of their earnings from user fees.</p> <p><b>Burkina Faso:</b> community-managed health centres established under the Bamako Initiative comprise one-third of public facilities and manage user fees (up to 10% of recurrent budget) for drugs mainly. Staff management is formally centralized. There are no clear accountability lines between community boards and health centre staff.</p> <p><b>Mali:</b> independent health centres are not-for-profit cooperative establishments owned, financed and managed by community associations. These health centres recruit their own staff. Few are as yet financially independent in practice.</p>
<p><b>General practitioners</b></p>	<p><b>Croatia:</b> previously centrally employed, salaried ambulatory care physicians. Now they are independent contractors.</p>

care. Tensions often occur when decision rights are not extended to managers (for example, when political pressure makes it impossible to dismiss staff), when accountability mechanisms are neither built into long-term contracts nor enforced through market discipline, and when the providers are not allowed to retain their surpluses or made responsible for their losses. The latter undermines the incentive to economize.

There is still considerable debate about whether long-term contracts with private providers create better incentives than similar contracts with public providers. Which incentives are most appropriate may depend on which goals have priority. The global trend is to try to avoid the inefficiencies and unresponsiveness that occur when a hierarchy becomes too rigid, while avoiding the opposite extreme of unregulated markets. The latter almost always undermine financial protection and may interfere with the strategic coordination needed to provide effective care.

## INTEGRATION OF PROVISION

As organizational units like hospitals or clinics become more autonomous, the service delivery system is at risk of becoming fragmented. Fragmentation may occur among similar provider configurations (hospitals, ambulatory clinics, or public health programmes) or between different levels of care. Such fragmentation has negative consequences for both the efficiency and the equity of the referral system unless explicit policies are introduced to ensure some sort of integration among the resulting semi-autonomous service delivery units.

When health services become fragmented, allocative efficiency suffers. For example, nonclinical health facilities designed to provide public health services in Poland and Hungary often engage in secondary prevention and a wide range of basic care because they are not adequately linked to ambulatory care networks. The university hospitals that have recently been made autonomous in Malaysia provide a wide range of inpatient and outpatient care for conditions that could have been treated effectively at lower levels in a community setting. The newly autonomous general practitioners in the Czech Republic have been quick to buy a large quantity of expensive equipment that is rarely used (59).

When organizational changes among providers cause fragmentation, disillusionment with a market-oriented system can lead to some vertical and horizontal reintegration, with more hierarchical control. Armenia, Hungary, New Zealand and the United Kingdom have recently experimented with such steps. Both the market model and the hierarchical model present problems; it is important not to forget the shortcomings of the centrally planned models that were apparent in countries as diverse as Costa Rica, Sri Lanka, Sweden, the United Kingdom and the former Soviet Union (59).

One way to preserve the virtues of autonomy for providers without fragmentation is via “virtual integration” instead of traditional vertical integration. Under vertical integration, a clinic takes orders from a hospital or a government department, limiting its responses to local needs. Virtual integration means using modern communication systems to share information quickly and without cumbersome controls. This is particularly valuable for referrals, and can include nongovernmental providers hard to incorporate under hierarchical schemes. Bangladesh and Ghana are experimenting with this innovation.

Even in the United States, vertical integration under health maintenance organizations (HMOs) is being eclipsed by virtual integration between the provider network HMOs,

other provider groups, and a globalized insurance industry. Vertical integration between production and distribution units is now being viewed as a coordination mechanism of last resort, and is used mainly when contractual alternatives are not available (60).

Efforts at virtual integration face three common problems, related to decentralization, separating purchasers from providers, and user charges. In many countries, there has recently been an increased enthusiasm for decentralization as a means of attaining a wide variety of policy and political goals in health as in other areas. The explicit objective of decentralization is often to improve responsiveness and incentive structures by transferring ownership, responsibility and accountability to lower levels of the public sector. This is usually done through a shift in ownership from the central government to local levels of the public sector – states or provinces, regions, districts, local communities, and individual publicly owned facilities.

A common difficulty with such reforms has been that the internal structural problems of the hospitals, clinics and public health facilities do not disappear during the transfer. In Uganda, decentralization did not close the financing gap experienced by many health facilities. In Sri Lanka, decentralization exposed weak management capacity but failed to address it. In Ghana, the unfunded social obligations were passed on to lower levels of government which did not have the financial capacity to absorb this responsibility because the proposed social insurance reforms had stalled. In many cases, central governments reassert control in a heavy-handed fashion when local governments deal with politically sensitive issues in a way that does not accord with the views of the national government on how such issues should be treated.

Where there is a split between purchasers and providers, similar tensions often arise. In Hungary and also in New Zealand there has been conflict between purchasing agencies situated in different branches of the government and still responsible for stewardship (such as ministries of health and finance) and the owners of the contracted providers (such as municipalities and local governments). In Hungary, constitutional powers were given to a self-governing National Health Insurance Fund that was controlled by the labour unions during the early 1990s. For about eight years, until the abolition of this arrangement in 1998, there was an open conflict between the Ministry of Finance and the Health Insurance Fund over fiscal policy and expenditure control. Providers were often not paid on time.

Finally, the introduction of user fees creates tensions between policy-based and prepaid purchasing and market-driven purchases of services by individual consumers. This has been especially true in many of the central Asian republics and in countries affected by the east Asia crisis, where the revenues channelled through policy-based purchasing have experienced a dramatic drop in recent years. This can undermine national policies on priority setting and cost containment, and as discussed in Chapter 2, it makes financing much less fair. The issue of how to organize purchasing as an integral part of the financing function is treated at more length in Chapter 5.

In order to attain the goals of good health, responsiveness and fair financial contribution, health systems need to determine some priorities and to find mechanisms that lead providers to implement them. This is not an easy task, because of two sources of complexity. Priorities should reflect a variety of criteria that are sometimes in conflict, and that requires a great deal of information that most health systems simply do not now have available. And to make priorities effective requires a mixture of rationing mechanisms, organizational structures, institutional arrangements and incentives for providers that must above all be consistent with one another and with the goals of the system.

## REFERENCES

1. **Beauchamp TL, Childress JF.** *Principles of biomedical ethics, 3rd ed.* Oxford, Oxford University Press, 1989.
2. **Gwatkin DR, Guillot M.** *The burden of disease among the global poor: current situation, future trends, and implications for strategy.* Washington, DC, and Geneva, The World Bank and the Global Forum for Health Research, 1999.
3. **Jha P, Chaloupka FJ, eds.** *Tobacco control policies in developing countries.* Oxford, Oxford University Press, 2000.
4. **Loudon IS, ed.** *Western medicine: an illustrated history.* Oxford, Oxford University Press, 1997.
5. **McKee M.** For debate – does health care save lives? *Croatian Medical Journal*, 1999, **40**(2): 123–128.
6. **Murray CJL et al.** *Development of WHO guidelines on generalized cost-effectiveness analysis.* Geneva, World Health Organization, 1999 (GPE Discussion Paper No. 4).
7. **Tengs TO.** *Dying too soon: how cost-effectiveness analysis can save lives.* Irvine, California, University of California, National Center for Policy Analysis, 1997 (Policy Report No. 204).
8. **Barnum H, Tarantola D, Setiady I.** Cost-effectiveness of an immunization programme in Indonesia. *Bulletin of the World Health Organization*, 1980, **58**(3): 499–503.
9. **Barnum H.** Cost savings from alternative treatments for tuberculosis. *Social Science and Medicine*, 1986, **23**(9): 847–850.
10. **Prescott N, de Ferranti D.** The analysis and assessment of health programs. *Social Science and Medicine*, 1985, **20**(12): 1235–1240.
11. **Mills A.** Economic evaluation of health programmes: application of the principles in developing countries. *World Health Statistics Quarterly*, 1985, **38**(4): 368–382.
12. **Mills A.** Survey and examples of economic evaluation of health programmes in developing countries. *World Health Statistics Quarterly*, 1985, **38**(4): 402–431.
13. **Lee M, Mills A, eds.** *Health economics in developing countries.* Oxford, Oxford University Press, 1984.
14. **World development report 1993 – Investing in health.** Washington, DC, The World Bank, 1993.
15. **Murray CJL, Lopez AD, eds.** *The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries and risk factors in 1990 and projected to 2020.* Cambridge, MA, Harvard School of Public Health on behalf of the World Health Organization and The World Bank, 1996 (Global Burden of Disease and Injury Series, Vol. I).
16. *The world health report 1999 – Making a difference.* Geneva, World Health Organization, 1999.
17. **Hammer J.** *Economic analysis for health projects.* Washington, DC, The World Bank, 1996 (Policy Research Working Papers, No. 1611).
18. **Jha P, Bangoura O, Ranson K.** The cost-effectiveness of forty health interventions in Guinea. *Health Policy and Planning*, 1998, **13**: 249–262.
19. **Peabody J et al.** *Policy and health: implications for development in Asia.* Cambridge, Cambridge University Press, 1999.
20. **Murray CJL, Kreuser J, Whang W.** Cost-effectiveness analysis and policy choices: investing in health systems. In: Murray CJL, Lopez A. *Global comparative assessments in the health sector: disease burden, expenditures and intervention packages.* Geneva, World Health Organization, 1994.
21. **Bobadilla JL.** Investigación sobre la determinación de prioridades en materia de salud [Inquiry into the determination of priorities in health]. In: Frenk J, ed. *Observatorio de salud.* Chapter 11. Mexico, Fundación Mexicana para la Salud, 1997 (in Spanish).
22. **Murray CJL, Lopez AD.** *Global health statistics.* Cambridge, MA, Harvard School of Public Health on behalf of the World Health Organization and The World Bank, 1996 (Global Burden of Disease and Injury Series, Vol. II).
23. **Wagstaff A.** *Inequalities in child mortality in the developing world. How large are they? How can they be reduced?* Washington, DC, The World Bank, 1999 (HNP Discussion Paper).
24. **Cotlear D.** *Peru: improving health care financing for the poor.* Washington, DC, The World Bank, 1999 (Country Studies).
25. **Wagstaff A.** *HNP module of Poverty Reduction Strategy Paper (PRSP) toolkit.* Washington, DC, The World Bank, 2000.
26. **Missialos E, Grand JL, eds.** *Health care and cost containment in the European Union.* Aldershot, Ashgate, 1999.
27. **Ham C.** Corporatization of UK hospitals. In: Preker AS, Harding A, eds. *Innovations in service delivery: the corporatization of hospitals.* Baltimore, Johns Hopkins University Press for The World Bank, 2000 (in press).

28. **Bobadilla JL 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.
29. **Bobadilla JL et al.** *El paquete universal de servicios esenciales de salud [Universal package of essential health services]*. Economía y salud, documentos para el análisis y la convergencia. Mexico, Fundación Mexicana para la Salud, 1994 (in Spanish).
30. **Robinson R.** Limits to rationality: economics, economists and priority setting. *Health Policy*, 1999, **49**: 13–26.
31. **Maynard A.** Rationing health care: an exploration. *Health Policy*, 1999, **49**: 5–11.
32. **Berryman SE et al.** *Assessing institutional capabilities*. Washington, DC, The World Bank, 1997 (World Bank Working Paper).
33. **Wilson JQ.** *Bureaucracy: what government agencies do and why they do it*. New York, Basic Books, 1989.
34. **Schmalensee R, Willing RD.** *Handbook of industrial organization, Vols I and II*. Amsterdam, Elsevier Science, 1989.
35. **Williamson O.** *The economic institutions of capitalism: firms, markets and relational contracting*. New York, Free Press, 1985.
36. **Williamson OE.** Comparative economic organization: the analysis of discrete structural alternatives. *Administrative Science Quarterly*, 1991, **36**(2): 269–296
37. **Barr N, ed.** *Labor markets and social policy in Central and Eastern Europe*. Oxford, Oxford University Press for The World Bank, 1994.
38. **World Bank.** *World development report 1996 – From plan to market*. New York, Oxford University Press, 1996: 123–132
39. **Lieberman S.** Corporatization of Indonesia hospitals. In: Preker AS, Harding A, eds. *Innovations in service delivery: the corporatization of hospitals*. Baltimore, Johns Hopkins University Press for The World Bank, 2000 (in press).
40. **Palmer N.** *The theory and evidence of contracting for primary care services: lessons for low and middle income countries*. London, London School of Hygiene and Tropical Medicine, 1999 (unpublished paper).
41. **Gregory R.** Accountability, responsibility, and corruption: managing the ‘public production process’. In: Boston J, ed. *The state under contract*. Wellington, New Zealand, Bridget Williams Books, 1995.
42. **Panzar J, Willing R.** *Economies of scale and economies of scope in multi-output production*. Murray Hill, NJ, Bell Laboratories, 1975 (Economic Discussion Paper No. 33).
43. **Teece DJ.** Economies of scope and the scope of the enterprise. *Bell Journal of Economic Behavior and Organization*, 1980, **1**: 223–247.
44. **Ho T.** Hospital restructuring in the European and Central Asian Regions. In: McKee M et al, eds. *The appropriate role of the hospital*. Copenhagen, World Health Organization Regional Office for Europe, 2000 (in press).
45. **Barnum H, Kutzin J.** *Public hospitals in developing countries: resource use, cost and financing*. Baltimore, Johns Hopkins University Press, 1993.
46. **Lipsky M.** *Street-level bureaucracy*. New York, Russell Sage, 1980.
47. **Tarimo E.** *Towards a healthy district: organizing and managing district health systems based on primary health care*. Geneva, World Health Organization, 1991.
48. **Conyers D, Cassels A, Janovsky K.** *Decentralization and health systems change: a framework for analysis*. Geneva, World Health Organization, 1993 (document WHO/SHS/NHP/93.2).
49. **Janovsky K, Travis P.** *Decentralization and health systems change: case study summaries*. Geneva, World Health Organization, 1998.
50. **Pressman J, Wildavsky A.** *Implementation*. Berkeley, University of California Press, 1973.
51. **Harding A, Preker AS.** Conceptual framework for organizational reform of hospitals. In: Preker AS, Harding A, eds. *Innovations in health service delivery: corporatization in the hospital sector*. Baltimore, Johns Hopkins University Press, 2000 (in press).
52. **Evans B.** General revenues financing: a comparison of US and Canada. In: Figueras J et al, eds. *Funding health care and long-term care: options in Europe*. Copenhagen, World Health Organization Regional Office for Europe, 2000 (in press).
53. **Robinson JC.** *The corporate practice of medicine: competition and innovation in health care*. Berkeley, University of California Press, 1999.
54. **Phua K.** Corporatization of hospitals in Singapore. In: Preker AS, Harding A, eds. *Innovations in health service delivery: corporatization in the hospital sector*. Baltimore, Johns Hopkins University Press, 2000 (in press).
55. **Scott G.** Corporatization in New Zealand hospitals. In: Preker AS, Harding A, eds. *Innovations in health service delivery: corporatization in the hospital sector*. Baltimore, Johns Hopkins University Press, 2000 (in press).

56. **Ham C.** Priority setting in health care: learning from international experience. *Health Policy*, 1997, **42**(1): 49–66.
57. **McKee M, Healy J.** *The role of the hospital in a changing environment*. European Observatory on Health Care Systems and London School of Hygiene and Tropical Medicine, 1999 (unpublished paper).
58. **Mills A.** Contractual relationship between government and the commercial private sector in developing countries. In: Bennett S, McPake B, Mills A. *Private health providers in developing countries: serving the public interest*. London, Zed Books, 1997.
59. **Preker AS, Feachem RGA.** *Market mechanisms and the health sector in Central and Eastern Europe*. Washington, DC, The World Bank, 1996 (Technical Paper No. 293).
60. **Klein B, Crawford RG, Alchian AA.** Vertical integration, appropriable rents, and the competitive contracting process. *Journal of Law and Economics*, 1978, **21**: 297–326.

## CHAPTER FOUR

# *What Resources are Needed?*

*Providing health care efficiently requires financial resources to be properly balanced among the many inputs used to deliver health services. Large numbers of physicians, nurses and other staff are useless without adequately built, equipped and supplied facilities. Available resources should be allocated both to investments in new skills, facilities and equipment, and to maintenance of the existing infrastructure. Moreover, these delicate balances must be maintained both over time and across different geographical areas. In practice, imbalances between investment and recurrent expenditures and among the different categories of inputs are frequent, and create barriers to satisfactory performance. New investment choices must be made carefully to reduce the risk of future imbalances, and the existing mix of inputs needs to be monitored on a regular basis. Clear policy guidance and incentives for purchasers and providers are necessary if they are to adopt efficient practices in response to health needs and expectations.*

## 4

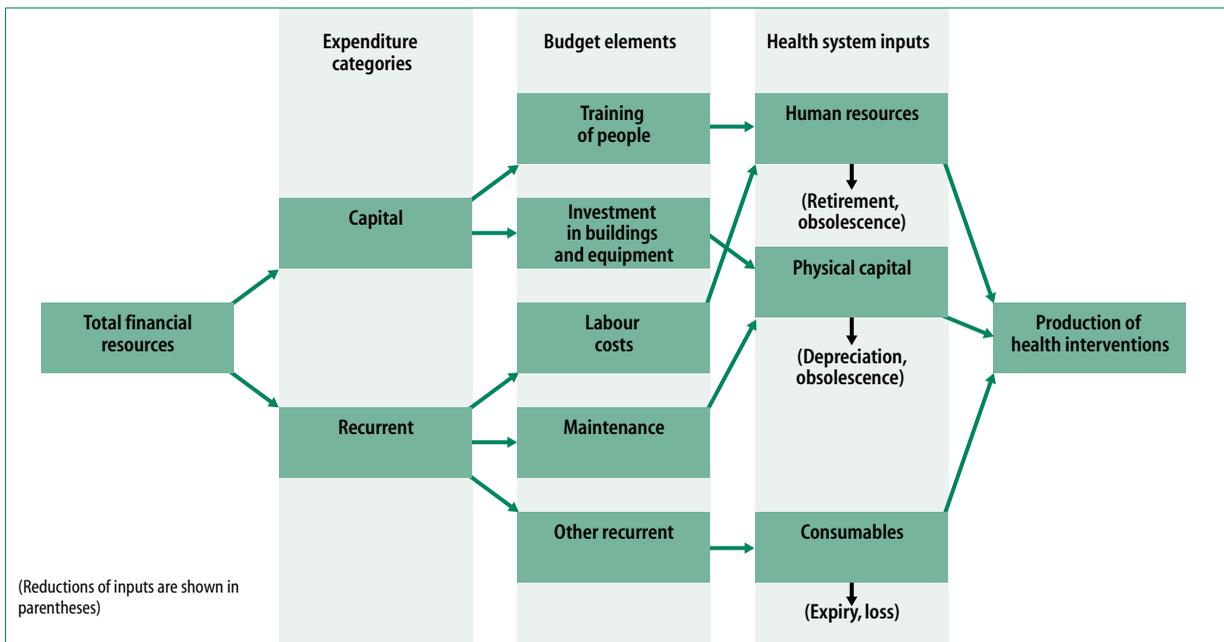
# WHAT RESOURCES ARE NEEDED?

## BALANCING THE MIX OF RESOURCES

The provision of health care involves putting together a considerable number of resource inputs to deliver an extraordinary array of different service outputs. Few, if any, manufacturing processes match the variety and rate of change of production possibilities in health. Figure 4.1 identifies three principal health system inputs: human resources, physical capital, and consumables. It also shows how the financial resources to purchase these inputs are of both a capital investment and a recurrent character. As in other industries, investment decisions in health are critical because they are generally irreversible: they commit large amounts of money to places and activities which are difficult, even impossible, to cancel, close or scale down.

The fact that some investment decisions lie outside the authority of the ministry of health makes the achievement of overall balance even more difficult. For example, the

**Figure 4.1 Health system inputs: from financial resources to health interventions**



training of doctors often comes under the ministry of education, and there may be private investment in facilities and equipment.

Capital is the existing stock of productive assets. Trained health workers and mobile clinics, as well as fixed assets, are part of the capital stock of the health system. Investment is any addition to this stock of capital, such as more pharmacists or additional vehicles. The typical productive lifetime of different investments will vary from as little as 1–2 years for certain equipment to 25–30 years or more for buildings and some kinds of professionals.

Technological progress influences the economic lifetime of a piece of capital: old investments quickly become outdated as new and improved technologies emerge. The way in which assets are managed also affects their lifetime. With proper handling and maintenance, buildings and vehicles lose their value more slowly. Without care and maintenance, health capital deteriorates rapidly. The planning of maintenance also needs to take the physical environment into account. For example, bad roads reduce the average lifetime of vehicles; so the planning of maintenance, operation and replacement of vehicles should allow for this.

Human capital can be treated conceptually in the same way as physical capital, with education and training as the key investment tools to adjust the human capital stock and determine the available knowledge and skills (1). Unlike material capital, knowledge does not deteriorate with use. But, like equipment, old skills become obsolete with the advent of new technologies, and human capital needs to be maintained too. Continuing education and on-the-job training are required to keep existing skills in line with technological progress and new knowledge. Human capital is also lost through retirement and death of individuals.

Investment also refers, in a broader sense, to any new programme, activity or project. Capital investment costs are all those costs that occur only once (to start up the activity), while the recurrent costs refer to the long-term financial commitment that usually follows from such an investment. If the available medical technology is seen as “capital”, and research and development as the investment tool to expand the technology frontier and develop new ideas, these concepts may also be applied to diagnostic equipment, medicines and the like.

Investment is the critical activity for adjusting capital stock and creating new and productive assets. Such adjustments typically occur gradually over time. Thus, the current physical infrastructure of hospital buildings and facilities in many countries is the product of an evolution that has taken many years. Among OECD countries, expenditures for investment in buildings and equipment are typically not more than 5% of total annual health care expenditures and are usually somewhat lower than they were 15 or 20 years ago: cost control has been enforced partly by controlling additions to capital.

In low income countries, however, there is greater variation. Investment levels can be substantially higher than the OECD figures, especially when physical infrastructure is being created or restored with the help of donor agencies. Countries such as Burkina Faso, Cambodia, Kenya, Mali and Mozambique report capital expenditures of between 40% and 50% of the total public health care budget in one or more years (2). A large percentage of the remaining recurrent budget usually pays for health care staff. This means that only a small fraction of the total budget is spent on the maintenance of physical and human capital and on consumable inputs, including pharmaceuticals. The balance between investments and other expenditures is more critical in low income countries as there is little room for mistakes. In general, however, very little is known about health investments in low income countries, even in the public sector. For the private sector, the available national health accounts estimates often have no data, or present implausibly high ratios of invest-

ment to total spending, maintained over many years. Not to know how much is being invested, and in what kinds of inputs, makes it nearly impossible to relate capital decisions to recurrent costs or assure that capital is not wasted or allowed to drain off funds needed for other inputs.

Even less data are available on the size of annual investments in education and training. These investment costs include medical and nursing schools, on-the-job training in different forms, and clinical research. Many players are involved and investments are often neither controlled by a single ministry nor guided by a common purpose. There is reason to believe that the sum of investments in human capital is far greater than investment in physical capital, at least in high income countries. As is the case for investment in physical capital, additions to human capital usually occur slowly over time. The training of a specialist, for example, can take 10 years or more of studies in medical school and on-the-job training. New investments in human capital also have long-term consequences, similar to investments in physical capital. The creation of a cadre of health workers with new skills, for example, will require a long-term investment in new curricula for basic and continuing education as well as a long-term commitment to paying their salaries.

## HUMAN RESOURCES ARE VITAL

Human resources, the different kinds of clinical and non-clinical staff who make each individual and public health intervention happen, are the most important of the health system's inputs. The performance of health care systems depends ultimately on the knowledge, skills and motivation of the people responsible for delivering services.

Furthermore, the human resources bill is usually the biggest single item in the recurrent budget for health. In many countries, two-thirds or more of the total recurrent expenditures reflect labour costs. But people would not be able to deliver services effectively without physical capital – hospitals and equipment – and consumables such as medicines, which play an important role in raising the productivity of human resources. Not only is a workable balance between overall health capital formation and recurrent activities needed, but the three input categories shown in Figure 4.1 should also be in equilibrium.

What treatment alternatives should be used for a certain illness or medical condition? Should services be offered at hospitals or primary care facilities? What is the level of skills and knowledge required to deliver this set of services? These questions have one thing in common. They are concerned with the degree of flexibility that exists in delivering health services, i.e. the possibility of substitution between one type of input and another, or the substitution of one form of care for another, all the while maintaining a constant level and quality of output. From a societal point of view, such positive substitution to achieve cost-effective delivery of services should be encouraged. A balanced combination of the different resource inputs will depend on identified health needs, social priorities and people's expectations.

Health systems are labour intensive and require qualified and experienced staff to function well. In addition to a balance between health workers and physical resources, there needs to be a balance between the different types of health promoters and care-givers. It would be an obvious waste of money to recruit physicians to carry out the simplest tasks. As a particular health system input is increased, the value added by each additional unit of input tends to fall (3). For example, where there are too few physicians, the arrival of another physician will have a positive effect on health care; but where there are already too many physicians, an additional physician is more likely to increase costs than improve care.

Some ways of dealing with imbalances among health care providers are outlined in Box 4.1.

A health system can have plentiful human resources, with excellent knowledge and skills, but still face impending crisis if future health needs, priorities and available resources are not taken into account. For example, where the education and training for junior doctors and nurses functions poorly, or where senior staff lack adequate time and resources to update their knowledge and skills, future shortfalls can be expected. Similarly, a health system with a skewed age distribution among staff towards the point of retirement poses a real problem. Thus, a health care system must balance investments in human capital to cover future needs as well as present demands. Some of the most critical and complex input problems relate to human resources (see Box 4.2).

Without functioning facilities, diagnostic equipment, and medicines, it does not matter if the knowledge, skills and staff levels are high. The delivery of services will still be poor. A lack of complementary inputs will also have a negative impact on staff motivation, a factor that influences the capacity of human resources. Motivation, however, depends not only on working conditions. Financial incentives and compensation, i.e. income and other benefits, are also important, as are the overall management of staff and the possibilities for professional advancement.

Inadequate pay and benefits together with poor working conditions – ranging from work in conflict zones to inadequate facilities and shortages of essential medicines and consumables – are frequently mentioned in less developed countries as the most pressing problems facing the health care workforce (4). In some countries, for example Bangladesh and Egypt, a clear majority of all publicly employed physicians see private paying patients to supplement income from their regular jobs. In Kazakhstan, “informal payments” are estimated to add 30% to the national health care bill (5). Possibilities for doctors to work privately in public institutions are being offered in some countries to neutralize an ongoing brain drain of qualified staff from the public sector. This strategy is considered successful in Bahrain, but experiences from Ghana and Nepal show that such incentives can lead to the diversion of scarce resources from public services and can induce professionals to engage in independent private practice (6).

People, as thinking creatures, are very different from machines and human capital cannot be managed in the same way as physical capital. First of all, human resources, and in

#### Box 4.1 Substitution among human resources

A large number of countries face an overall shortage of physicians. Other countries that are following a long-term strategy to shift resources to primary care find that they have too many specialists and too few general practitioners. Many are dealing with the problems by substituting among various health care-givers.

*Reorientation of specialist physi-*

*cians.* While limiting admissions to specialist training and changing internship programmes is a long-term strategy to balance the professional distribution of physicians, the reorientation of specialists into family practice is a short-run substitution strategy being used, for example, in central and eastern Europe.

*Substitution for other health professionals.* The training of a physician

may cost three times more than that of a nurse.<sup>1</sup> As a result, training of more nurses as well as other health professionals may be a cost-effective substitute for physicians. In Botswana, training of more nurse practitioners and pharmacists has offset the lack of physicians in some areas.<sup>2</sup>

*Introduction of new cadres.* Ensuring a closer match between skills

and function may demand the creation of new cadres. In Nepal, an educational programme allowed health assistants and other health workers in rural areas to train for higher professional postings.<sup>3</sup>

<sup>1</sup> World development report 1993 – Investing in health. New York, Oxford University Press for The World Bank, 1993.

<sup>2</sup> Egger D, Lipson D, Adams O. Achieving the right balance: the role of policy-making processes in managing human resources for health problems. Geneva, World Health Organization, 2000 (Issues in health services delivery, Discussion paper No. 2, document WHO/EIP/OSD/2000.2).

<sup>3</sup> Hicks V, Adams O. The effects of economic and policy incentives on provider practice. Summary of country case studies using the WHO framework. Geneva, World Health Organization, 2000 (Issues in health services delivery, Discussion paper No. 5, document WHO/EIP/OSD/2000.8 (in press)).

particular physicians, determine the use of other available inputs. An oversupply of physicians will almost certainly mean an oversupply of the kind of services that physicians provide. The high density of private physicians working in urban areas of many middle income countries, such as Thailand, usually correlates with frequent use of expensive equipment and laboratory testing, and with more services of sometimes doubtful value being provided to the urban population. In Egypt, the high ratio of physicians – for every occupied bed in Egypt there are two physicians – combined with extensive self-medication explain the very high use of drugs. According to estimates, the poorest households in Egypt spend over 5% of their income on drugs alone (2).

Incentives and management related to human resources have an indirect impact on the use of other resources as well. For example, many payment systems provide physicians and providers with incentives to use more or less medical equipment, laboratory testing and medicines. In Bangladesh, physicians get 30–40% of the laboratory charges for each referral generated, creating a clear interest to expand the volume of such services (2). In both China and Japan, many physicians derive part of their income from the sale of drugs which they prescribe. In many countries, the use of branded drugs instead of generics is still common, and this can to a large extent be blamed on the incentives offered to physicians and pharmacists by pharmaceutical producers. Lack of the skills needed to assess technology and control quality is an additional factor causing imbalances among resources.

Another difference between human and physical capital, which affects how people are managed, is that physicians, nurses and other health workers are not motivated only by present working conditions, income and management. They are also influenced by what they believe those conditions will be in the future, based on past experiences, views expressed by others and current trends. If qualified staff believe that future payment, benefits and working conditions will deteriorate, their job-related decisions and motivation will reflect that belief. This “shadow of the future” can easily result in a continuing negative spiral towards lower motivation and performance.

A first step to prevent such a development is to find a sustainable balance among the different types of resources and between investment and recurrent costs. Perhaps the most

#### Box 4.2 Human resources problems in service delivery

*Numerical imbalances.* A recent study of human resources in 18 low and middle income countries, one or more in each of the WHO regions, indicates that most countries experience varying degrees of shortages in qualified health personnel. In sub-Saharan Africa in particular, the limited training capacity and low pay for qualified health workers causes severe problems in service delivery. Elsewhere, for example in Egypt, oversupply is a problem. Generally, shortages and oversupply are defined relative to countries in the same region and at similar levels of development. Oversupply, thus,

may be absolute, as is the case for specialist physicians in many countries of eastern Europe and central Asia, or relative to geographical location.

*Training and skill mix imbalances.* Health care workers are often unqualified for the tasks they perform because of a shortage of training opportunities, as in many African countries, or a mismatch between available skills and the needs and priorities of the health care system, as in eastern Europe and central Asia. The number of physicians and other health personnel with a certain type of training or qualification, however, tells only part of the story.

Neither formal training nor professional affiliation necessarily equates with skill in dealing with specific problems.

*Distribution imbalances.* Almost all countries have some urban/rural imbalances among their human resources and face problems in meeting the needs of specific groups such as poor or handicapped people or ethnic minorities. It is almost universally true that providers tend to concentrate in urban areas. In Cambodia, 85% of the population live in rural areas, but only 13% of the government health workers work there. In Angola, 65% live in rural areas, but 85% of health pro-

professionals work in urban areas. In Nepal, only 20% of rural physician posts are filled, compared to 96% in urban areas.

*Failure of past public policy approaches.* Although progress has been made in recent years to develop national policies and plans for human resources for health, they are not fully implemented in most countries. Moreover, very few countries monitor and evaluate the progress and impact of policy implementation.

important part of such a balance is to ensure that there are individual incentives to invest in human capital in the form of improved earnings, career opportunities and working conditions. Indeed, many low and middle income countries have increased pay or benefits as a key strategy for developing human resources and improving delivery of services to meet health needs and priorities (7). Public sector pay in Uganda rose by 900% (in nominal terms) between 1990 and 1999, which represents a doubling in real terms (8).

In general there are no easy answers in the area of human resources development. Left unmanaged, human skills markets take years, even decades, to respond to market signals.

And, unlike physical capital, human resources cannot be scrapped when their skills are no longer needed or obsolete; even laying off public sector health workers is often so difficult that it can only be achieved as part of a broader policy to reform the civil service.

Public intervention to produce the required balance is thus essential to reduce waste and accelerate adjustment. Some successful experiences are summarized below but many problems remain (7).

*Utilization levels, mix and distribution.* The relative prices of different skill categories should guide decisions about their most efficient mix, where labour markets are functioning. There are no absolute norms regarding the right ratio of physicians or nurses to population; rules of thumb are often used. Generally, shortages or oversupply are assessed on the basis of need and priorities combined with comparisons with neighbouring countries or those at a similar level of development. Such assessment requires sound data about available human resources and their geographical and professional distribution: such information is often lacking. In Guinea-Bissau, 700 "ghost" workers were removed from the payroll of the Ministry of Finance, following an inventory of the health care workforce. Cambodia's 1993 survey of health workers revealed a poorly distributed and largely unregistered workforce, with widely differing competencies (2).

Three types of human resource strategy have been pursued with some success:

- making more efficient use of available personnel through better geographical distribution;
- greater use of multiskilled personnel where appropriate;
- ensuring a closer match between skills and functions.

The latter strategy responds to a widespread problem. Formal training of health workers, particularly for more highly skilled staff, too seldom reflects the actual tasks being performed. This is both wasteful and demoralizing.

Some success has been recorded with mandatory service and multiple incentives (financial, professional, educational, etc.) to make otherwise unattractive technical or geographical areas more appealing, as has been done in Canada and the Scandinavian countries to deploy staff in their northern regions. Countries such as Fiji, Oman and Saudi Arabia have successfully recruited foreign workers to fill critical gaps, as an interim strategy. This strategy can, however, create other difficulties and tensions. Oman at present has a policy to recruit primarily a domestic workforce, as the pool of potential medical students has increased.

*Intake training and continuing education.* A clear case can be made for strong public sector involvement in training and in monitoring the quality of continuing education to stimulate the development of human resources in targeted areas. New public health schools have recently been established in Hungary and Jamaica to meet needs for professionals with skills in epidemiology, statistics, management and health education. They aim to integrate initial formal training, subsequent continuing education, and actual service provision.

This has two potential benefits. It ensures that training has strong practical foundations, and it continually exposes service providers to new thinking and development. In countries with large rural populations several strategies have been used to recruit staff to rural areas. Examples are intake of medical students from rural areas and training in the locations where physicians will later practise.

A related problem concerns the brain drain of trained staff from low income countries to wealthier countries or from the public sector to the private sector within a country. The more successful trainees often emigrate, tempted by higher standards of practice and living abroad. Many Jamaican nurses have migrated to the United States. Physicians migrate from Egypt and India to other countries in the Middle East and to the USA and Europe. Inadequate pay and benefits rank as the most serious problem confronting the public sector health workforce in many countries, with growing formal and informal private practice as a consequence. Service contracts that require a certain number of years in public service, especially when the training is state sponsored, have been implemented in the Philippines and the United Republic of Tanzania, and are common in Latin America but there are attendant difficulties. The staff concerned are usually junior, placements are short term and unpopular, mentoring arrangements are seldom adequate, and overall geographical imbalance is little affected. Globalization has led to greater mobility of staff and opportunity for overseas training, and students who qualify abroad may wish to stay in the country where they were trained.

## ADJUSTING TO ADVANCES IN KNOWLEDGE AND TECHNOLOGY

Growth in the available knowledge or advances in technology – such as new drugs or diagnostic equipment – can substantially increase the capacity of human resources to solve health problems, and thereby improve the performance of a health care system. New knowledge is also a challenge to each country's existing input balance, as relative prices change and the efficient mix of resources alters (9). In the past few decades, revolutionary advances in medicine and technology have shifted the boundaries between hospitals, primary health care, and community care (10). Corresponding resource shifts in health systems have been much slower to emerge.

Antibiotic drugs provide one example of new knowledge affecting cost structures. Since their introduction in the 1940s, patients suffering from a bacterial infection have most often been cared for at home or at outpatient clinics rather than in special hospitals, significantly reducing costs and improving outcomes. The recent growth of unregulated self-treatment and the increasing incidence of drug-resistant bacteria have compromised some of these gains. There is now a need for active stewardship to regulate the quality of diagnosis, prescribing and compliance. Vaccines have similarly altered the strategy and costs of tackling epidemic diseases such as measles and poliomyelitis, and new vaccines will continue to necessitate re-thinking to ensure an efficient mix of inputs in national health strategy.

All countries – rich as well as poor – need to find and maintain a reasonable balance between inputs. The choices involved in finding this balance, however, vary depending on the amount of total resources available. In a poor country, the possibilities of investing in modern medical technologies or paying for modern medicines are very limited. Moving from the use of essential drugs to new and expensive drugs for cardiovascular diseases would mean an enormous opportunity loss in terms of health outcome for a poor country. This difference in opportunities across countries also has an impact on the optimal balance

between resources (see Box 4.3).

Some input prices are determined locally; others are set in international markets. In most countries, prices for human resources (incomes for physicians, nurses and other health care personnel) are determined nationally, and the general income level for each country or region will be an important determinant. Prices for such items as patented drugs and medical equipment, on the other hand, are determined in a global market. Although differences in income levels across countries will induce manufacturers and distributors of medicines and equipment to differentiate prices somewhat, stewards of individual country health systems are far less able to influence these prices than the prices of human resources. International stewardship is needed to represent the interests of consumers in low income countries that face heavy burdens of infectious and parasitic diseases. This type of stewardship, led by agencies such as WHO and the World Bank, will assume increasing importance as globalization of the economy continues and free trade agreements are implemented.

## PUBLIC AND PRIVATE PRODUCTION OF RESOURCES

With the exception of skilled human resources, most inputs used for health services are produced in the private sector, with varying degrees of public stewardship over the level and mix of production, distribution, and quality. For example, local markets successfully produce most consumables and unskilled labour. Government intervention is needed mainly to ensure that quality and safety standards are met, that reliable information is available about the products, and that a fair competitive environment exists.

Other inputs, such as manufactured pharmaceuticals and specialized medical equipment, often face barriers to entry into the market in the form of patents and licensing requirements, manufacturing standards, large initial investment costs, expensive research, and long development periods. This gives the manufacturers of these inputs considerable market power to abuse by manipulating prices and demand. Strong policy measures are therefore needed, such as anti-trust legislation, limited formularies, generic drug policies, bulk purchasing, and formal technology assessments (11–13). Furthermore, by procuring

### Box 4.3 A widening gap in technology use?

A vast quantity of valuable medical technologies and innovative clinical methods have been developed over the past decades and many more are on the way. Unfortunately, the new possibilities are not open to all because of the lack of available income in some countries. Diseases that are treated effectively in rich countries by professional staff using modern technology are handled by unskilled staff or informally at home in less developed countries. Moreover, some

of these diseases are more prevalent in the poorest countries.

Medicines are now available for *HIV/AIDS* that can, at a huge cost, at least postpone further development of the disease. But treatment patterns and resource inputs for *HIV/AIDS* currently follow different paths in different countries. In poor countries, *HIV/AIDS* is still a disease without treatment alternatives. The sick are mainly taken care of informally at home or in institutions with predominantly unskilled staff. South Africa has improved the availability

of *HIV* treatment by obliging insurers to cover its cost for members of insurance schemes.

*Malaria* transmission can be prevented by means of house spraying, insecticide-treated nets, chloroquine prophylaxis, and so on, but such measures are not always available to the people who need them most. Several different projects to develop a *malaria* vaccine are under way.<sup>1</sup> A breakthrough in this research would present a tremendous opportunity to improve quality of life and prevent death. Such a tech-

nological breakthrough would also demand a new mix of resources, but only for those countries that could afford the new vaccine.

For *tuberculosis*, the incidence of bacterial resistance to first-line drugs is increasing. It is of major concern, for example, in the Russian Federation. Lack of effective medical treatment and improper use of medicines continue to create obstacles to dealing with this escalating problem.<sup>2</sup>

<sup>1</sup> *The world health report 1999 – Making a difference*. Geneva, World Health Organization, 1999.

<sup>2</sup> *Global tuberculosis control: WHO report 2000*. Geneva, World Health Organization, 2000 (document WHO/CDS/TB/2000.275).

medicines and medical technologies on the international market, countries can ensure that local producers remain competitive (14, 15).

Publicly subsidized production of consumables, pharmaceuticals and medical equipment often leads to low quality, lack of innovation, outmoded technology, inefficient production modalities and distribution delays. The most striking example of this occurred in the former Soviet Union. Most countries that have followed this model have quickly fallen behind in productivity and production technology. Many Western firms that entered the pharmaceutical and medical equipment market in central and eastern Europe during the early 1990s found it cheaper and easier to build new factories than to convert and modernize the old capital stock (16–18).

Decisions on physical capital, such as hospitals and other large facilities, require more public attention. Ambulatory clinics, laboratories, pharmacies, cottage hospitals, and other small clinical facilities often have small capital requirements, and private providers may be able to finance these themselves or through small personal loans in parallel to public investments. In the case of large hospitals, most countries have in the past relied heavily on public investments. Investment decisions in this area have consequences that may last for 30–40 years or more. Once built, hospitals are politically difficult to close. The need for strong public policies, however, does not necessarily mean the public financing of the entire capital stock. Increasingly, many countries are looking to the private sector to support investments in their health system even when the resulting facilities will not have for-profit objectives, and the running costs will be publicly financed (19). Chapter 6 illustrates some pitfalls of developing joint venture investments, and the different skills required for compe-

#### Box 4.4 The Global Alliance for Vaccines and Immunization (GAVI)

Every year, nearly three million children die from diseases that could be prevented with currently available vaccines, yet nearly 30 million of the 130 million children born every year are not receiving vaccinations of any kind. The great majority of unreached children – 25 million – live in countries that have less than US\$ 1000 per capita GNP.

The Global Alliance for Vaccines and Immunization (GAVI) is a coalition of public and private interests that was formed in 1999 to ensure that every child is protected against vaccine-preventable diseases. GAVI partners include national governments, the Bill and Melinda Gates Children's Vaccine Program, the International Federation of Pharmaceutical Manufacturers Associations (IFPMA), research and technical health institutions, the Rockefeller Foundation, UNICEF, the World Bank Group, and WHO.

GAVI is seeking to close the growing gap of vaccine availability between industrialized and developing countries. Beyond the six basic vaccines of the Expanded Programme on Immunization (against poliomyelitis, diphtheria, whooping cough, tetanus, measles and tuberculosis), newer vaccines, such as those for hepatitis B, *Haemophilus influenzae* type b (Hib), and yellow fever are now widely used in industrialized countries. A major priority is to see that all countries of the world achieve at least 80% immunization coverage by 2005. Based on current assumptions of vaccine delivery costs it is estimated that an additional \$226 million annually are needed to reach this level of coverage in the poorest countries with the traditional EPI vaccines; to cover the same number of children with the newer vaccines, according to the guidelines adopted at GAVI's first board meeting, would require an additional \$352 million.

At the second meeting of the GAVI board, held during the World Economic Forum in Davos in February 2000, the GAVI partners discussed policies for attaining the 80% immunization objective and announced a multimillion-dollar global fund for children's vaccines. Governments, businesses, private philanthropists, and international organizations are working together to manage these resources so as to provide the protection of immunization to children in all countries, under the campaign title of "The Children's Challenge". Members of GAVI argue that protecting the world's children against preventable diseases is not only a moral imperative but an essential cornerstone of a healthy, stable global society.

All countries with incomes of less than \$1000 per capita GNP (74 countries worldwide, with the majority in Africa) have been invited to express their interest in collaborating with GAVI in this campaign.

Nearly 50 countries, from all WHO regions, have already provided details of their immunization activities and needs. Resources from the fund will primarily be used to purchase vaccines for hepatitis B, *Haemophilus influenzae* type b (Hib), and yellow fever, and safe injection materials.

It is envisaged that GAVI partners at the country level will collaborate with national governments to help close the gaps identified in the country proposals other than those directly related to the provision of vaccines. By placing more of the responsibility for providing the necessary information and commitment on the countries themselves, the GAVI partners are hoping that resulting efforts will be more country-driven and therefore more sustainable.

tent stewardship of such developments. With regard to the training of specialized labour and the generation of knowledge, the story is similar. There is a need for strong public involvement in setting the policy agenda and ensuring adequate regulation, but private capital can be mobilized to support investments in both training and research activities.

The dominant force underlying the 20th-century revolution in health services has been the new global knowledge made possible by research and development. Chapter 1 echoes *The world health report 1999* in arguing that today's health systems have a clear responsibility to provide the knowledge for the health systems of tomorrow (20). Investment in knowledge which can be used by all has a special merit (see Box 4.4). Although most research and development is, and should be, financed through private capital, there should be public involvement in supporting such endeavours and directing them towards areas of greatest need. Attempts to directly manage the dynamics of research and development from the top, however, often fail. Experience suggests that indirect approaches and providing the research community with appropriate incentives will be more successful. Once again, imaginative international stewardship may make a vital difference.

## THE LEGACY OF PAST INVESTMENTS

Past investments in the poorest countries in the world have focused on the accumulation of physical infrastructure. Such programmes have often been supported by multilateral and bilateral donor agencies in the hope that they would lead to improved performance and that the countries themselves would be able to collect sufficient public money to cover recurrent costs (21). In reality, resources to maintain and operate both physical and human capital have often been insufficient. Health facilities are unable to function well because of poor maintenance and shortages of essential drugs and supplies. Vehicles are often immobile for lack of repair and maintenance. For example, in Ghana at one point in 1992, 70% of Ministry of Health vehicles were reported immobile, pending repair at government workshops. Reorganization of maintenance and repair arrangements and budget practice led to rapid improvement, but Ghana's recent experience is widespread. Even in places where vehicles are mobile, fuel is often a scarce resource. These are just some examples of imbalances that all lead to reduced performance, a shorter lifespan of the physical infrastructure and low staff morale. In terms of physical capital, the situation is often irreversible. The cost of renovating is higher than the cost of building anew.

Lack of necessary skills, poor cost information systems, rigid budgeting systems, and fragmentation of tasks – such as separation of responsibility for investment from operating budgets – are further reasons behind input imbalances. If information on needed quantities is not available, it is difficult to estimate reasonable budget levels for inputs such as consumables and fuel for transportation. If providers are then responsible for holding each of these budget lines, serious barriers are created to delivering health services effectively. Shortages of essential production inputs too frequently coexist with unused funds in over-compartmentalized health budgets.

Chapter 6 considers the recent development of formal partnerships, such as sector-wide approaches (SWAPs), between government and groups of donors. On the capital investment front, donors could do much better. External agencies have contributed to unbalanced input mixes by focusing on highly visible investments without adequate consideration of compatibility with other investments (for example, with respect to spare parts), or recurrent costs. Political success in health system investment is seldom the ally of long-term

sustainability. There are often incentives in less developed countries for decision-makers to accept donor support irrespective of the long-term consequences on the balance among existing resources or between investments and recurrent costs. For example, Sri Lanka accepted a donor contribution of a 1000-bed hospital: to operate it took needed resources away from many other activities. Competing agendas among donors have led to further fragmentation in responsibility and short-term thinking (22).

## HEALTH CARE RESOURCE PROFILES

Large differences in the mix of resources used by high and low income countries can partly be explained by differences in relative prices. A full system of national health accounts offers the most complete information on health system inputs and their prices, as discussed below. In a poor country, unskilled human resources will be relatively cheap, whereas medical technology, facilities and highly qualified staff will be expensive. As a result, a large percentage of the total public budget is often allocated to investment. Once staff have been paid from the recurrent budget, there will be little left to spend on equipment, medicines, consumables and maintenance of facilities. This is evidence that there are simply too many staff, often reflecting the training of staff relative to population norms or need-based planning, rather than in accordance with resource-based planning. Drug consumption in low income countries – often high in relative terms – is to a large extent financed privately by consumers through out-of-pocket payment. In a more developed country, spending on consumables will be much higher in absolute numbers, but still low in relative terms due to the fact that human resources will be more expensive. But even for countries with comparable income levels there are sometimes wide differences in country-specific resource profiles.

Figure 4.2 shows resource profiles for four high income countries: Denmark, Sweden, the United Kingdom, and the United States (23). Each country's input level, on each of eight inputs, is expressed as a percentage of the highest value for that indicator in the group: the figures do not show "best performance" in the sense of Chapter 2 but simply compare input levels.

The United States is at or close to the maximum on every input. On expenditure and technology it is at the maximum of this group of countries. Sweden has the largest stock of human resources and beds and, with Denmark, the highest drug spending. The United Kingdom is within the boundary set by this group on every input: it is particularly far from the maximum on expenditure per capita and technology (magnetic resonance imaging (MRI) and computerized tomography (CT) scanners). However, in terms of beds and drugs the UK is comparable to the rest of the group and higher than the USA.

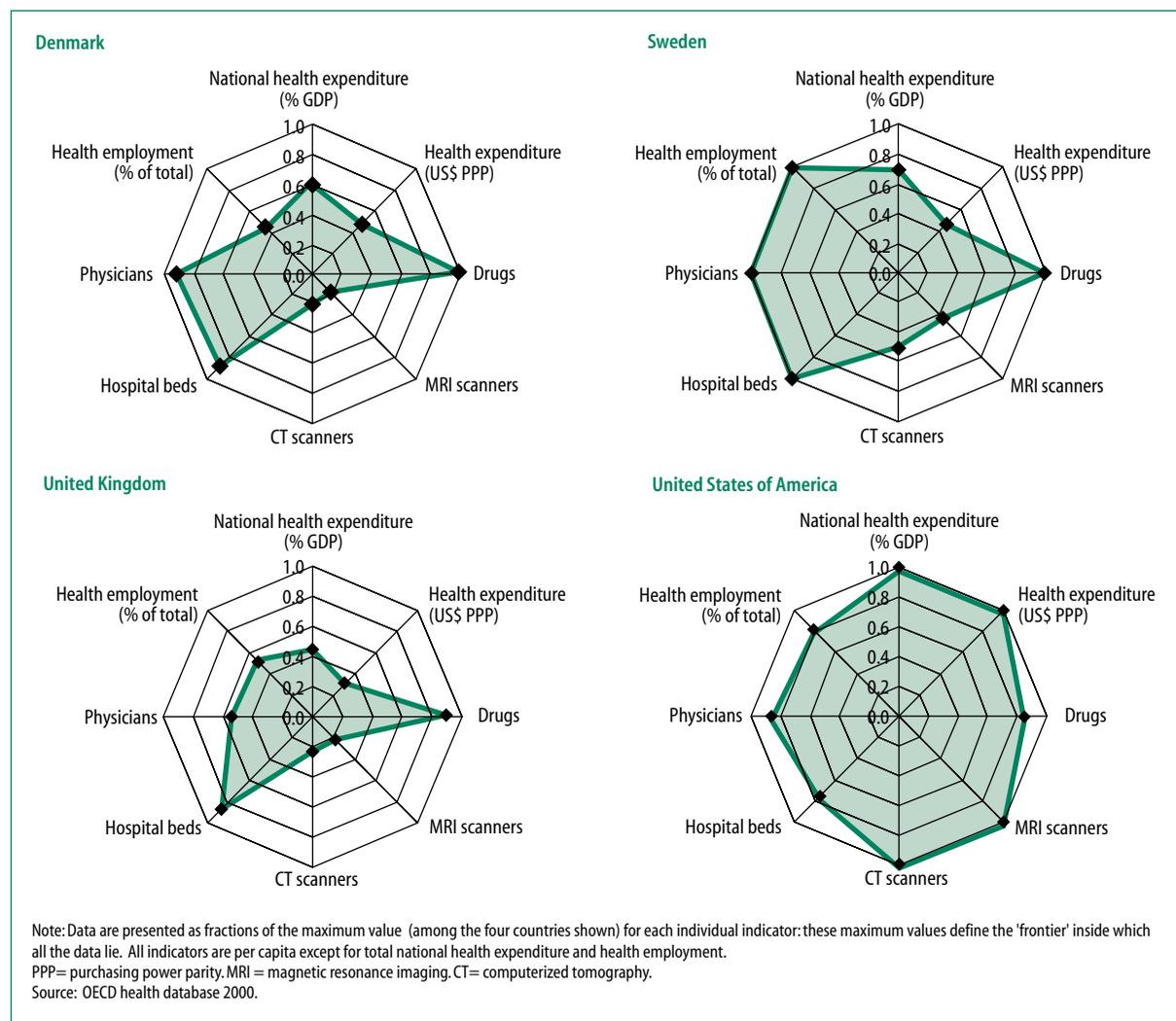
This simple comparison between countries shows clear differences in terms of input mix. The differences can in part be explained by past conditions of competition and payment methods among US hospitals, which have focused more on quality than on price and cost-effectiveness. Relative price differences also play a role. The 'medical arms race' in Sweden and particularly in the UK has been more restrained under global budgets. This also means that the US health system (and ultimately US tax-payers and health insurance payers) pays a larger share of the global costs involved in bringing new medical technologies and medicines to the market. Trend data (not in the diagrams) show that, with respect to MRI and CT scanners, both Sweden and the UK are catching up with levels in the United States. This supports the view that the US health system is an early adopter of new

medical technology. The relative price of physicians and nurses in Sweden is low compared to that in the United States, and the different input mixes illustrate a degree of substitutability between human resources and other health inputs.

Figure 4.3 shows similar resource profiles for Egypt, Mexico, South Africa and Thailand. These four middle income countries spend substantially less on all types of health care resources than the group of high income countries. As is the case for the group of high income countries, there are considerable contrasts in the mix of resources and these differences do not seem to be due primarily to differences in income or prices.

South Africa is at the maximum of the group for expenditure, nurses, beds and MRI scanners, while it is furthest from the maximum for drugs and physicians (with Thailand). Egypt has the lowest total health expenditure per capita within the group, but the highest ratio of physicians and the second highest level of drug consumption. Both physicians and drugs in Egypt are mostly paid for directly by patients out of pocket. Some 80% of physicians' income is estimated to come from private practice, and households finance close to 60% of total drug costs through direct payments (2). Doubts have been raised about skill

Figure 4.2 Health systems input mix: comparison of four high income countries, around 1997

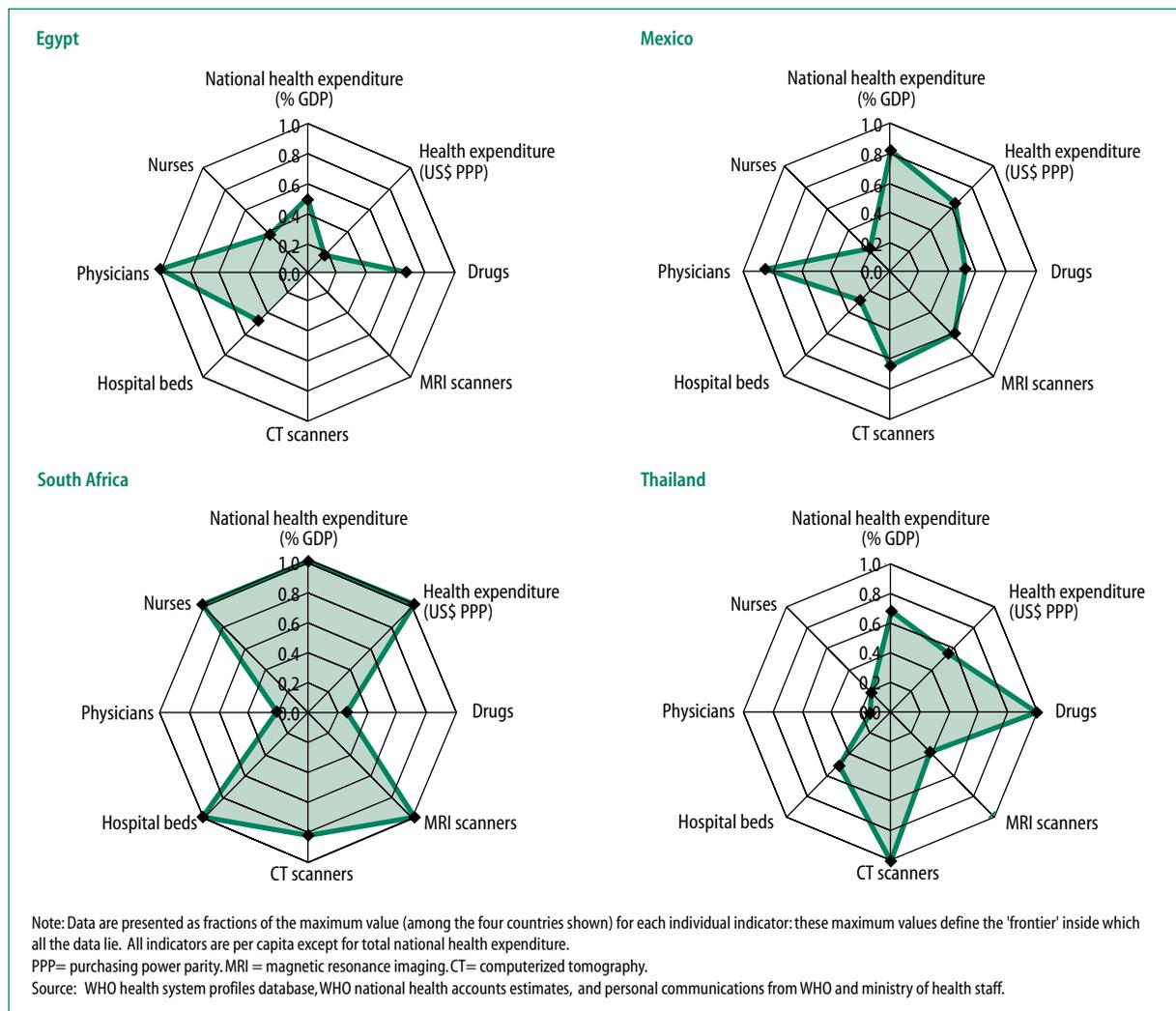


levels of physicians. There is extensive use of branded as opposed to generic drugs. In part this pattern of drug use is explained by little knowledge and poor perception of generic drugs by consumers, combined with extensive self-medication. Irrational prescribing by physicians and dispensing by pharmacists of expensive drugs are other important explanatory factors.

Mexico has a high ratio of physicians and, together with Thailand, the lowest ratio of nurses within the group. It is estimated that about 15% of all physicians in Mexico are either inactive, underemployed or unemployed (2). Despite this evidence of surplus, there are a large number of unfilled posts in rural areas. In contrast, Thailand and South Africa have a low ratio of physicians. Indeed, Thailand's health authorities estimate that at least another 10 000 physicians are needed (2). The ratios of physicians to nurses show great contrasts: in South Africa nurses greatly outnumber physicians, possibly due to greater international mobility of doctors, but in Egypt and Mexico the proportions are reversed.

The distribution of available resources between urban and rural areas is a major problem in all four middle income countries but is not illustrated by the figure, which presents

Figure 4.3 Health systems input mix: comparison of four middle income countries, around 1997



only averages. Physicians mostly work where health status levels are highest. The distribution of resources across ethnic groups is a particular problem in South Africa. In Thailand, most of the high technology equipment is concentrated in urban hospitals, whereas the use of technology at the primary care level is scarce. Most of the about 900 physicians produced annually in Thailand remain in urban areas, and shortages of qualified staff in rural areas are expected to persist.

## CHANGING INVESTMENT PATTERNS

Experience points to political difficulties in changing existing investment patterns and resource profiles. Every euro, bhat or kwacha spent on health service delivery or investment is income to someone and therefore creates a vested interest (24). If the income is large, this “someone” will lobby for more resources and resist changes that do not match his or her particular interests. Such lobbying and resistance come from both the medical industry and from labour groups. Attempts to reform systems for paying providers are often highly controversial, as are changes in medical school admissions or educational programmes. Lobbying also comes from interest groups and politicians. Health care investments usually attract popular support and it can be difficult to rearrange investments in favour of a new balance. This will often be the case even if large imbalances exist compared to social priorities. Vested interests and lobbying related to the distribution of cost and benefits are important factors in the inertia that has to be overcome in order to change the existing capital structure and mix of resource inputs.

The predominant investment emphasis in the health system over past decades has been on hospitals and specialist care. In addition to the other forces opposing primary health care, discussed in Chapter 1, investment decisions played a part. The allocation of investment capital to hospital buildings is not the main reason. More importantly, the focus on specialist care entailed investments in the employment and training of human resources to staff hospitals. The focus on hospital care led to a rapid accumulation of beds. In high income countries, the accumulation of beds was accompanied and driven by rapid technological change, resulting in a greater intensity of care and increasing costs. Population ageing, and the accompanying higher health systems utilization rates by elderly people, maintained this upward pressure on bed supply.

In less developed countries the accumulation of beds has been accompanied by much slower technological change and slower cost increases, but also by less intensity of care, inadequate maintenance of facilities and lower quality of services because of a lack of funding for recurrent costs (24). In many middle and low income countries, occupancy rates at public hospitals have been low. In Mexico, for example, occupancy rates at Ministry of Health hospitals have been 50% on average, because of inadequate staffing and maintenance, with consequent inefficient use of existing resources (25).

During the 1990s, many countries started to reduce the size of their hospital sector and many small hospitals, in particular, were closed or used for other purposes (see Box 4.5). Reflecting technical progress and lower costs in ambulatory care, the number of beds has declined and the average length of stay has been reduced. Closure of small hospitals and emergency wards and a declining number of beds mean that new strategies will have to be developed to respond to fluctuating demand, with greater integration among providers, transport of patients, pooling of resources and information as key components.

With fluctuating demand, there is a need for some surplus capacity to absorb changes in demand. The influenza epidemic that swept across Europe in 1999–2000 revealed that sur-

plus bed-capacity to deal with sudden changes in health need is limited in many countries, for example in the UK (26). In other industries (electricity supply, public transport) temporary surges in demand can be met through peak-load pricing. As discussed more fully in Chapter 3, rationing by price is not an acceptable allocation mechanism in the health system. Excess demand that cannot be repressed by higher prices must be accommodated by other means of rationing – postponing non-emergency care, transferring patients, shortening inpatient stays and so on.

In countries with a hierarchical planning structure, resource allocation and investment planning is often incremental. Last year's budget is often the starting point for next year's planning. Ongoing activities are usually not questioned: this greatly reduces the country's scope for shifting to a more cost-effective overall allocation of inputs. The planning process can be described as a game in which ministries that deliver services (such as ministries of health) call for increased resources, while the guardians of the treasury try to maintain expenditure at its previous level (27). This approach is attractive because of its simplicity but it demands growing budgets. If budgets are declining, departments should really scrutinize the full range of ongoing programmes and activities, prioritizing activities for possible cutting or elimination. Public bureaucracies typically try to maintain the status quo by cutting costs across the board without changing overall priorities, and without taking special account of the need to protect targeted geographical areas or sub-groups of the population.

Agencies will try, for as long as possible, to maintain what they judge to be critical expenditure such as salaries and cut down on expenditure that does not immediately damage health system performance (22). Planned investments are delayed and ongoing constructions are left incomplete. In the hope that financial crises are short-lived, health systems may decrease their spending on long-term investments in human and physical capital and even on recurrent costs for maintenance, medicines and other consumables. This will eventually constrain severely the capacity of human capital and health system performance. Investments, by their nature, tend to be more volatile than recurrent expenditures: they occur in discrete chunks and then require smaller but regular operating expenditures. Short-

#### Box 4.5 Investment in hospitals in countries of the former Soviet Union prior to policy reform

The majority of health care resources in the former Soviet Union were controlled from the top by ministries of health. Central government managed investments and the consequent accumulation of resources in physical and human capital. The structure of service delivery was determined by such norms as beds or physicians per thousand inhabitants. The result was high hospital capacity.

In the early 1990s the bed ratio for most of the former Soviet Union was considerably higher than that in many western European

countries. The number of physicians per capita, most of them allocated to the hospital sector, was also high in comparison with many western countries. Indications of inefficiency were given by long lengths of stay and moderate occupancy rates, especially in small hospitals. The effectiveness of hospital services was also influenced adversely by the poor quality of facilities and medical equipment. Many of the small district hospitals had no more than 4–5 m<sup>2</sup> per bed, and some of the smallest hospitals had no radiology services, and inadequate heating or water.

For example, a 1989 survey found that 20% of Russian hospitals did not have piped hot water, 3% did not even have piped cold water, and 17% lacked adequate sanitation facilities. The survey also found that every seventh hospital and polyclinic needed basic reconstruction. A similar survey of facilities in 1988 found substantial underinvestment in maintenance of polyclinics and hospitals, with 19% of polyclinics and 23% of hospitals rated as either being in a "disastrous" condition or requiring full reconstruction.

In the 1990s, reductions in the number and use of hospitals were

an essential part of reforms. A combination of overcapacity and poor quality of physical resources had become a major distortion in the input mix of these countries. Where facilities were not closed, or used as nursing homes or for other functions, they were upgraded and used more effectively in the referral system. In many cases, however, changes have been modest because of political difficulties in transferring resources from one use to another.

term postponement or cutting of investment may be an appropriate response to a crisis, but it requires an overall picture of capital and recurrent resources as well as a likely time horizon for the crisis. Without these, ad hoc chopping of planned investments will create imbalances and inefficiency.

## THE WAY FORWARD

Clear symptoms of imbalances between resources include poor performance, deteriorating facilities, and low working morale among staff. Often, skilled human capital moves to the private sector or to wealthier regions. Physical capital deteriorates in a more visible way. The patient turns to the private sector in search of better quality care.

Whatever a country's income level, there exist efficient ways to allocate health system inputs that will allow the health system to function at its best. The efficient mix will vary over time and across countries, depending on relative prices among inputs, country specific health needs and social priorities. In less developed countries, setting priorities will surely be much harder, and the balance between investments and recurrent costs more critical. Health care systems face major challenges when there is a rapid change either in technology or in available financial resources as a result of a turbulent macroeconomic environment. The failure of health care decision-makers to respond to such a shift in conditions will lead to suboptimal health system performance.

For very different reasons, both developed and less developed countries record imbalances between the available inputs. Because of the rapid technological changes in health services, imbalances have been the rule rather than the exception in developed countries. Problems are much more visible in less developed countries, where imbalances have often been caused by lack of management skills and a decline in the available financial resources. Although some imbalances are likely to exist even in well-functioning health systems, much more could be done to correct them rapidly or even prevent them.

A first step is to create a general awareness of the problem by documenting the various resources used and the performance of health systems. Sound data on the existing numbers and distribution of human resources, especially when linked to data on health system performance, can also contribute to the formulation of policies and plans to address problems. Figures 4.2 and 4.3 present fragments of information on inputs. But they do not reveal how input mix affects performance.

National health accounts (NHAs) offer a more comprehensive framework for bringing together data on inputs and for communicating with various stakeholders on future investment policies. NHAs give a broad picture, which enables ministries to lead health care services through reforms and difficult times. When Finland's economy went into crisis with the loss of its export market with the Russian Federation in the early 1990s, health policy-makers were able to use their NHA information to restore productivity in the health system.

More appropriate cost information and accounting systems would also make it easier to achieve a balance among inputs, for example by establishing more reliable budget estimates. A general awareness and improved information through NHAs and accounting systems will not result in any change, however, without a parallel and widespread commit-

ment by health care decision-makers to address the fundamental problems.

Such commitment is best supported by a combination of stewardship – oversight and influence – and more scope for decentralized decision-making by purchasers and providers. Central authority over major investment decisions is essential. This does not mean that all such decisions need to be made centrally. But central policy and guidance, through a bidding or certification process, are necessary to ensure overall coordination between public and private investment decisions, and with the recurrent funding capacity of the public sector. The worst mistake is to promote or allow investments when their running costs cannot be met. Central policy on drugs and major technology registration, the development of essential drugs lists and treatment guidelines, quality assurance and bulk tendering will continue to be necessary. Purchasers and providers need incentives and opportunities to challenge the prevailing allocation of inputs in order to discover the best way to respond to health needs, social priorities and expectations. Rigid hierarchical approaches to balancing resources usually result in reactive rather than continuous change; shortages of essential inputs on one hand and unspent funds on the other are likely to be common problems. But decentralized decision-making among providers must be controlled and guided through active purchasing and appropriate payment mechanisms to meet overall priorities. Decentralized decision-making on the details of service and intervention arrangements also requires new strategies for human resources and investments in planning and management skills at all levels.

Without such explicit stewardship of all input sources and monitoring of developments, there will be too much discretion among decentralized units to engage in opportunistic behaviour. Such behaviour, either at central or decentralized levels, will also deter donor agencies from supporting decentralization, for example through sector-wide approaches and common funding pools. It will also be a reason for aid recipients to mistrust attempts to bring about donor coordination (22). Decentralization does not mean a lack of accountability in resource management, nor that central government should opt out of planning and monitoring. It should be designed to increase accountability and should give central government and ministries a new role, focusing on overall regulation and monitoring.

As part of that new role, the impact of new medical technologies should be assessed and regulatory practice developed in consultation with the important stakeholders. Such assessment of new technologies requires documentation on existing practice and use of resources. This further emphasizes the importance of monitoring. For the less developed countries, donor agencies need to take existing and possible imbalances into account when drawing up support strategies. The information base provided by a consistent use of NHAs will provide a good starting point for a common understanding of existing imbalances.

In both rigid hierarchical systems and in decentralized systems without accountability, proper incentives and stewardship, imbalances among resources will be more difficult to correct and prevent. Such imbalances often create huge problems in their own right, but they will also induce further problems by giving wrong signals to the health care labour market and the industry that supports health services. Well-performing, cost-effective health systems that respond to health needs based on explicit priorities will give both the medical industry and medical schools the incentives to invest properly in research and development, in educational programmes and in the physical inputs essential to the production of better health.

## REFERENCES

1. **Becker GS.** *Human capital. A theoretical and empirical analysis with special reference to education*, 3rd edition. Chicago, The University of Chicago Press, 1993.
2. WHO health system profiles database.
3. **Folland S, Goodman AC, Stano M.** *The economics of health and health care*. New York, Macmillan Publishing Company, 1993.
4. **Berckmans P.** *Initial evaluation of human resources for health in 40 African countries*. Geneva, World Health Organization, Department of Organization of Health Services Delivery, 1999 (forthcoming document).
5. **Ensor T, Savelyeva L.** Informal payments for health care in the Former Soviet Union: some evidence from Kazakhstan. *Health Policy and Planning*, 1999, **13**(1): 41–49.
6. **Hicks V, Adams O.** *The effects of economic and policy incentives on provider practice. Summary of country case studies using the WHO framework*. Geneva, World Health Organization, 2000 (Issues in health services delivery, Discussion paper No. 5, document WHO/EIP/OSD/2000.8, in press).
7. **Egger D, Lipson D, Adams O.** *Achieving the right balance: the role of policy-making processes in managing human resources for health problems*. Geneva, World Health Organization, 2000 (Issues in health services delivery, Discussion paper No. 2, document WHO/EIP/OSD/2000.2).
8. *Public sector pay reform project, final report*. Ministry of Public Service, Accra, Uganda, 1999, WHO health system profiles database.
9. **Weisbrod BA.** The health care quadrilemma: an essay on technological change, insurance, quality of care, and cost containment. *Journal of Economic Literature*, 1991, **24**: 523–552.
10. *European health care reform. Analysis of current strategies*. Copenhagen, Denmark, World Health Organization, 1997 (European Series No. 72).
11. *Action Programme on Essential Drugs*. Geneva, World Health Organization, 1999.
12. *Public-private roles in the pharmaceutical sector: implications for equitable access and rational use*. Geneva, World Health Organization, 1997 (Health economics and drugs, DAP Series No. 5).
13. **Banta HD, Luce BR.** *Health care technology and its assessment: an international perspective*. Oxford, Oxford University Press, 1993.
14. *Medical equipment procurement manual*. Washington, DC, The World Bank, 1998.
15. *Standard bidding documents: procurement of health sector goods*. Washington, DC, The World Bank, 2000.
16. **Van Gruting CWD ed.** *Medical devices: international perspective, Part VIII – Geographical situations in central and east Europe*. Amsterdam, Elsevier, 1994.
17. **Issakov A, Richter N, Tabakow S.** Health care equipment and clinical engineering in central and eastern Europe. *New World Health*, 1994: 167–171.
18. *Russian Federation medical equipment project*. Washington, DC, The World Bank, 1996 (Report No. 14968-RU).
19. *Private hospital study*. Washington, DC, International Finance Corporation, 1998.
20. *The world health report 1999 – Making a difference*. Geneva, World Health Organization, 1999.
21. **Lee K.** Symptoms, causes and proposed solutions. In: Abel-Smith B, Creese A, eds. *Recurrent costs in the health sector: problems and policy options in three countries*. Geneva, World Health Organization, 1989 (document WHO/SHS/NHP/89.8).
22. **Walt G et al.** Health sector development: from aid coordination to resource management. *Health Policy and Planning*, 1999, **14**(3): 207–218.
23. **Anell A, Willis M.** International comparison of health care resources using resource profiles. *Bulletin of the World Health Organization*, 2000, **78** (in press).
24. **Reinhardt U.** *Accountable health care: is it compatible with social solidarity?* London, The Office of Health Economics, 1997.
25. **Barnum H, Kutzin J.** *Public hospitals in developing countries: resource use, cost and financing*. Baltimore, John Hopkins University Press, 1993.
26. The health service: bedridden. *The Economist*, 15 January 2000: 31–34.
27. **Wildavsky A.** *Budgeting: a comparative theory of budgetary processes*. Boston, Little, Brown & Co., 1975.

## CHAPTER FIVE

# *Who Pays for Health Systems?*

*Choices for financing health services have an impact on how fairly the burden of payment is distributed. Can the rich and healthy subsidize the poor and sick? In order to ensure fairness and financial risk protection, there should be a high level of prepayment; risk should be spread (through cross-subsidies from low to high health risk); the poor should be subsidized (through cross-subsidies from high to low income); the fragmentation of pools or funds should be avoided; and there should be strategic purchasing to improve health system outcomes and responsiveness.*

## 5

## WHO PAYS FOR HEALTH SYSTEMS?

## HOW FINANCING WORKS

*H* health care expenditures have risen from 3% of world GDP in 1948 to 7.9% in 1997. This dramatic increase in spending worldwide has prompted societies everywhere to look for health financing arrangements which ensure that people are not denied access to care because they cannot afford it. Providing such access to all citizens has long been a cornerstone of modern health financing systems in many countries. The main function of the health system is to provide health services to the population, and this chapter concentrates on health financing as a key to effective interaction between providers and citizens. It discusses the purpose of health financing, and the links between health financing and service delivery, through purchasing. The factors affecting the performance of health financing are also examined.

The purpose of health financing is to make funding available, as well as to set the right financial incentives for providers, to ensure that all individuals have access to effective public health and personal health care. This means reducing or eliminating the possibility that an individual will be unable to pay for such care, or will be impoverished as a result of trying to do so.

To ensure that individuals have access to health services, three interrelated functions of health system financing are crucial: revenue collection, pooling of resources, and purchasing of interventions. The main challenges are to put in place the necessary technical, organizational and institutional arrangements so that such interactions will protect people financially the fairest way possible, and to set incentives for providers that will motivate them to increase health and improve the responsiveness of the system. The three functions are often integrated in a single organization, and this is currently the case in many health systems in the world. Although this chapter discusses the three functions separately, it does not imply that an attempt should be made to separate them in different organizations. There is, however, an increasing trend to introduce a separation between financing and provision.

*Revenue collection* is the process by which the health system receives money from households and organizations or companies, as well as from donors. Contributions by donors are discussed in Box 5.1. Health systems have various ways of collecting revenue, such as general taxation, mandated social health insurance contributions (usually salary-related and almost never risk-related), voluntary private health insurance contributions (usually risk-related), out-of-pocket payment and donations. Most high income countries rely heavily

### Box 5.1 The importance of donor contributions in revenue collection and purchasing in developing countries

Donor contributions, as a source of revenue for the health system, are of key importance for some developing countries. The absolute amounts of such aid have been large in recent years in Angola, Bangladesh, Ecuador, India, Indonesia, Mozambique, Papua New Guinea, the United Republic of Tanzania and several eastern European countries, but in the larger countries aid is usually only a small share of total health spending or even of government expenditure. In contrast, several

countries, particularly in Africa, depend on donors for a large share of total expenditure on health. The fraction can be as high as 40% (Uganda in 1993) or even 84% (Gambia in 1994) and exceeds 20% in 1996 or 1997 in Eritrea, Kenya, The Lao People's Democratic Republic and Mali. Bolivia, Nicaragua, the United Republic of Tanzania and Zimbabwe have obtained 10% to 20% of their resources for health from donors in one or more recent years.

Most aid comes in the form of projects, which are separately developed and negotiated between each donor and the national authorities. Although by no means unsuccessful, international cooperation through projects can lead to fragmentation and duplication of effort, particularly when many donors are involved, each focusing on their own geographical or programme priorities. Such an approach forces national authorities to devote significant amounts of time and effort to dealing with donors' priorities

and procedures, rather than concentrating on strategic stewardship and health programme implementation. Donors and governments are increasingly seeing the need to move away from a project approach towards wider programme support to long-term strategic development that is integrated into the budgetary process of the country. In this respect, sector-wide approaches have been effective in countries such as Bangladesh, Ghana and Pakistan.<sup>1</sup>

<sup>1</sup> Cassels A, Janovsky K. Better health in developing countries: are sector-wide approaches the way of the future? *The Lancet*, 1998, 352:1777–1779.

on either general taxation or mandated social health insurance contributions. In contrast, low income countries depend far more on out-of-pocket financing: in 60% of countries at incomes below \$1000 per capita, out-of-pocket spending is 40% or more of the total whereas only 30% of middle and high income countries depend so heavily on this kind of financing (see Table 5.1).

In most social insurance and voluntary private insurance schemes, revenue collection and pooling are integrated in one organization and one purchasing process. For organizations relying mainly on general taxation, such as ministries of health, collecting is done by the ministry of finance and allocation to the ministry of health occurs through the government budgetary process.

*Pooling* is the accumulation and management of revenues in such a way as to ensure that the risk of having to pay for health care is borne by all the members of the pool and not by each contributor individually. Pooling is traditionally known as the “insurance function” within the health system, whether the insurance is explicit (people knowingly subscribe to a scheme) or implicit (as with tax revenues). Its main purpose is to share the financial risk associated with health interventions for which the need is uncertain. In this way it differs from collecting, which may allow individuals to continue bearing their own risks from their own pockets or savings. When people pay entirely out of pocket, no pooling occurs.

**Table 5.1 Estimated out-of-pocket share in health spending by income level, 1997**  
(number of countries in each income and expenditure class)

Estimated annual per capita income (US\$ at exchange rate)	Estimated share in total expenditure on health (%)						Total
	Under 20	20–29	30–39	40–49	50–59	60 and over	
Under 1000	7	10	9	7	11	19	<b>63</b>
1000–9999	16	18	23	15	8	8	<b>88</b>
10 000 and over	19	7	4	5	0	2	<b>37</b>
All income classes	42	35	36	27	19	29	<b>188</b>

Source: WHO national health accounts estimates: income unknown for three countries.

For public health activities and even for aspects of personal health care – such as health check-ups – for which there is no uncertainty or the cost is low, funds can go directly from collecting to purchasing. This is an important consideration with regard to the regulation of mandatory pooling schemes, as consumer preferences for insurance packages often focus on interventions of high probability and low cost (relative to the household capacity to pay), although these are best paid for out of current income or through direct public subsidies for the poor.

Pooling reduces uncertainty for both citizens and providers. By increasing and stabilizing demand and the flow of funds, pooling can increase the likelihood that patients will be able to afford services and that a higher volume of services will justify new provider investments.

*Purchasing* is the process by which pooled funds are paid to providers in order to deliver a specified or unspecified set of health interventions. Purchasing can be performed passively or strategically. Passive purchasing implies following a predetermined budget or simply paying bills when presented. Strategic purchasing involves a continuous search for the best ways to maximize health system performance by deciding which interventions should be purchased, how, and from whom. This means actively choosing interventions in order to achieve the best performance, both for individuals and the population as a whole, by means of selective contracting and incentive schemes. Purchasing uses different instruments for paying providers, including budgeting. Recently, many countries, including Chile (1, 2), Hungary (3), New Zealand (4, 5), and the United Kingdom (6–8), have tried to introduce an active purchasing role within their public health systems.

## PREPAYMENT AND COLLECTION

Traditionally, most policy discussions regarding health system financing centre around the impact of public versus private financing on health system performance. Chapter 3 clarifies the central role of public financing in public health. For personal health care, however, it is not the public–private dichotomy that is most important in determining health system performance but the difference between prepayment and out-of-pocket spending. Thus, private financing, particularly in developing countries, is largely synonymous with out-of-pocket spending or with contributions to small, voluntary and often highly fragmented pools. In contrast, public or mandatory private financing (from general taxation or from contributions to social security) is always associated with prepayment and large pools. The way policy-makers organize public financing or influence private financing will affect four key determinants of health system financing performance: the level of prepayment; the degree of spreading of risk; the extent to which the poor are subsidized; and strategic purchasing.

A health system where individuals have to pay out of their own pockets for a substantial part of the cost of health services at the moment of seeking treatment clearly restricts access to only those who can afford it, and is likely to exclude the poorest members of society (9–12). Some important health interventions would not be financed at all if people had to pay for them, as is the case for the public good type of interventions discussed in Chapter 3 (13). Fairness of financial risk protection requires the highest possible degree of separation between contributions and utilization. This is particularly so for interventions that are high cost relative to the household's capacity to pay.

In addition to affording protection against having to pay out of pocket and, as a result, facing barriers to access, prepayment makes it possible to spread the financial risk among

members of a pool, as discussed later in the chapter. Individual out-of-pocket financing does not allow the risk to be shared in that way. In other words, as already proposed by *The world health report 1999* (14), there has to be prepayment for effective access to high-cost personal care.

The level of prepayment is mainly determined by the predominant revenue collection mechanism in the system. General taxation allows for maximum separation between contributions and utilization, while out-of-pocket payment represents no separation. Why then is the latter so generally used, particularly in developing countries? (15).

The answer is that separation of contributions from utilization requires the agencies responsible for collection to have very strong institutional and organizational capacity. These attributes are lacking in many developing countries. Thus, although the highest possible level of prepayment is desirable, it is usually very difficult to attain in low income settings where institutions are weak. Relying on prepaid arrangements, particularly general taxation, is institutionally very demanding. General taxation, as the main source of health financing, demands an excellent tax or contribution collecting capacity. This is usually associated with a largely formal economy, whereas in developing countries the informal sector is often predominant. While general taxation on average accounts for more than 40% of GDP in OECD countries, it accounts for less than 20% in low income countries.

All other prepayment mechanisms, including social security contributions and voluntary insurance premiums, are easier to collect, as the benefit of participating is linked to actual contributions. In most cases, participation in social insurance schemes is restricted to formal sector workers who contribute through salary deductions at the workplace. This makes it easier for the social security organization to identify them, collect contributions and possibly exclude them from benefits if no contribution is made. Similarly, identification and collection is easier for voluntary health insurance and community pooling arrangements. Nevertheless, such prepayment still requires large organizational and institutional capacity compared to out-of-pocket financing.

In developing countries, therefore, the objective is to create the conditions for revenue collecting mechanisms that will increasingly allow for separation of contributions from utilization. In low income countries, where there are usually high levels of out-of-pocket expenditure on health and where organizational and institutional capacity are too weak to make it viable to rely mainly on general taxation to finance health, this means promoting job-based contribution systems where possible, and facilitating the creation of community or provider-based prepayment schemes. Evidence shows (16, 17), however, that although the latter are an improvement over out-of-pocket financing, they are difficult to sustain and should be considered only as a transition towards higher levels of pooling or as instruments to improve the targeting of public subsidies in health. In middle income countries, with more formal economies, strategies to increase prepayment as well as pooling arrangements include strengthening and expanding mandatory salary-based or risk-based contribution systems, as well as increasing the share of public financing, particularly for the poor.

Although prepayment is a cornerstone of fair health system financing, some direct contribution at the moment of utilization may be required in low income countries or settings to increase revenues where prepayment capacity is inadequate. It can also be required in the form of co-payment for specific interventions with a view to reducing demand. Such an approach should only be used where there is clear evidence of unjustified over-utilization of the specific intervention as a result of prepayment schemes (moral hazard). The use of co-payment has the effect of *rationing* the use of a specific intervention but does not have the effect of *rationalizing* its demand by consumers. When confronted with co-payments,

people, particularly the poor, will reduce the amount of services demanded (even to the extent of not demanding a service at all) but will not necessarily be more rational in distinguishing when to demand services or which services they need to demand. Therefore, using user charges indiscriminately will indiscriminately reduce demand, hurting the poor in particular.

Free-of-charge services do not translate automatically into unjustified over-utilization of services. Services that are free of direct charge are in reality not necessarily free or affordable, particularly for the poor, because of other costs associated with seeking health care, such as the cost of medication (when not available free of charge), under-the-table payments, transportation, or time lost from work (18, 19).

Given its potentially negative impact on necessary services, especially for the poor, co-payment should not be chosen as a source of financing except for low-cost relatively predictable needs. Rather, it can be used as an instrument to control over-utilization of specific interventions (when such over-utilization is evident) or to implicitly exclude services from a benefit package when explicit exclusion is not possible. Because of the desirability of separating contributions from utilization, out-of-pocket payment should not be used unless no other alternative is available. All prepaid arrangements are preferable, except for low-cost interventions for which the administrative costs involved in prepayment arrangements might not be worthwhile.

## SPREADING RISK AND SUBSIDIZING THE POOR: POOLING OF RESOURCES

Pooling is the main way to spread risks among participants. Even when there is a high degree of separation between contributions and utilization, prepayment alone does not guarantee fair financing if it is on an individual basis only – that is, via medical savings accounts. Individuals would then have limited access to services after their savings were exhausted. It is claimed for medical savings accounts, which have been implemented in Singapore and in the United States, that they reduce moral hazard and give consumers the incentive to buy services more rationally, but while there is evidence of reduced expenditure and of substantial savings among those who receive tax benefits and can afford to save (20), there is no evidence of more rational purchasing. And individual financing fosters fee-for-service payment and makes it harder to regulate the quality of provision (21). People with a high risk of having to use services, such as the sick and the elderly, would be denied access because they could not save enough from their income. On the other hand, the healthy and the young, whose risk is usually low, might prepay for a long time without needing the services for which they had saved. In this case, mechanisms allowing for cross-subsidies from the young and healthy to the sick and old would benefit the former without damaging the latter. Thus, systems as well as people benefit from mechanisms that not only increase the degree of prepayment for health services, but also spread the financial risk among their members.

Although prepayment and pooling are a significant improvement over purely out-of-pocket financing, they do not take questions of income into account. As a result of large pools, society takes advantage of economies of scale, the law of large numbers, and cross-subsidies from low-risk to high-risk individuals. Pooling by itself allows for equalization of contributions among members of the pool regardless of their financial risk associated with service utilization. But it also allows the low-risk poor to subsidize the high-risk rich. Societies interested in equity are not indifferent to who is subsidized by whom. Therefore, health

financing, in addition to ensuring cross-subsidies from low to high risk (which will happen in any pool, unless contributions are risk-related), should also ensure that such subsidies are not regressive (see Figure 5.1).

Health systems throughout the world attempt to spread risk and subsidize the poor through various combinations of organizational and technical arrangements (22). Both risk- and income-related cross-subsidies could occur among the members of the same pool, for example in single pool systems such as the Costa Rican social security organization and the national health service in the UK, or via government subsidies to a single or multiple pool arrangement.

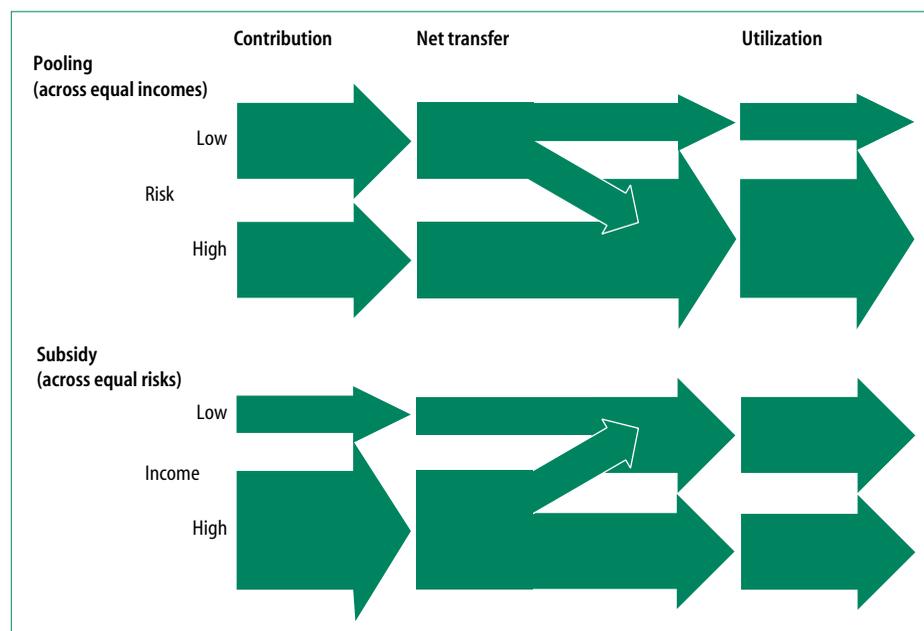
In practice, in the majority of health systems, risk and income cross-subsidization occurs via a combination of two approaches: pooling and government subsidy. Cross-subsidization can also occur among members of different pools (in a multiple pool system) via explicit risk and income equalization mechanisms, such as those being used in the social security systems of Argentina (23), Colombia (24) and the Netherlands (25). In these countries, the existence of multiple pools allows members of pools to have different risk and income profiles. Without some compensatory mechanisms, such arrangements would offer incentives for pooling organizations to select low risks, and to exclude the poor and the sick.

Even under single pool organizations, decentralization, unless accompanied by equalization mechanisms for resource allocation, may result in significant risk and income differences among decentralized regions. Brazil has introduced compensatory mechanisms in the allocation of revenues from the central government to the states to reduce such differences (26).

Table 5.2 shows four country examples of different arrangements for spreading risk and subsidizing the poor. Some organizational arrangements are less efficient than others in ensuring that these two objectives are achieved, particularly if the arrangements facilitate fragmentation, creating numerous small pools. Collecting, pooling, purchasing and provi-

**Figure 5.1 Pooling to redistribute risk, and cross-subsidy for greater equity**

(arrows indicate flow of funds)



sion imply flows of funds from sources to providers through a variety of organizations, which may perform only one, or several of these tasks. Figure 5.2 illustrates the structure of health system financing in four countries which differ greatly in the degree to which there is formal pooling of funds and purchasing, rather than consumers paying directly to providers without any sharing of risks.

Large pools are better than small ones because they can increase resource availability for health services. The larger the pool, the bigger the share of contributions that can be allocated exclusively to health services. A large pool can take advantage of economies of scale in administration and reduce the level of the contributions required to protect against uncertain needs, while still ensuring that there are sufficient funds to pay for services. Given that needs vary unpredictably, the estimation for an individual could be unaffordably large. By reducing this uncertainty, the pool is able to reduce the amount set aside as a financial reserve to deal with variations in the health expenditure estimates for its members. It can then use the funds released for more and better services.

Fragmentation of the pool – in other words, the existence of too many small organizations involved in revenue collection, pooling and purchasing – damages performance of all three tasks, particularly pooling. In fragmented systems, it is not the number of existing pools and purchasers that matters, but that many of them are too small. In Argentina, prior to the 1996 reforms, there were more than 300 pooling organizations (*Obras Sociales Nacionales*) for formal sector workers and their families, some with no more than 50 000 members. The administrative capacity and financial reserves required to ensure financial viability for the small ones, together with the low wages of their beneficiaries, guaranteed that their benefit packages were very limited. A similar problem occurs with community

**Table 5.2 Approaches to spreading risk and subsidizing the poor: country cases**

Country	System	Spreading risk	Subsidizing the poor
<b>Colombia</b>	Multiple pools: multiple competing social security organizations, municipal health systems and Ministry of Health.	Intra-pool via non-risk-related contribution and inter-pool via a central risk equalization fund. Mandated minimum benefit package for all members of all pools.	Intra-pool and inter-pool: salary-related contribution plus explicit subsidy paid to the insurer for the poor to join social security; supply side subsidy via the Ministry of Health and municipal systems.
<b>Netherlands</b>	Multiple pools: predominantly private competing social insurance organizations.	Intra-pool via non-risk-related contribution and inter-pool via central risk equalization fund.	Via risk equalization fund, excluding the rich.
<b>Republic of Korea</b>	Two main pools: national health insurance and the Ministry of Health.  National health insurance, however, only covers 30% of total health expenditures of any member.	Intra-pool via non-risk-related contribution.  Explicit single benefit package for all members.	Salary-related contribution plus supply side subsidy via the Ministry of Health and national health insurance from Ministry of Finance allocations.  Public subsidy for insurance for the poor and farmers.
<b>Zambia</b>	Single predominant formal pool: Ministry of Health/Central Board of Health.	Intra-pool, implicit single benefit package for all in the Ministry of Health System and at state level. Financed via general taxes.	Intra-pool via general taxation. Supply side subsidy via the Ministry of Health.

Figure 5.2 Structure of health system financing and provision in four countries

<b>Bangladesh (1996/97)</b>					
Revenue collection	General taxation	Donors	Out-of-pocket		Other
Pooling	Ministry of health		Other governmental	No pooling	
Purchasing	Ministry of health		Other governmental	Individual purchasing	
Provision	Ministry of health		Private providers		

<b>Chile (1991–1997)</b>				
Revenue collection	General taxation		Social insurance	Out-of-pocket
Pooling	Public health insurance fund (FONASA)		Private insurance (ISAPREs)	No pooling
Purchasing	Public health insurance fund (FONASA)		Private insurance (ISAPREs)	Individual purchasing
Provision	Other governmental	National health service	Private providers	

<b>Egypt (1994/95)</b>				
Revenue collection	General taxation	Donors	Social insurance	Out-of-pocket
Pooling	Ministry of health	Other governmental	Social insurance	No pooling
Purchasing	Ministry of health	Other governmental	Social insurance	Individual purchasing
Provision	Ministry of health	Ministry of health	Other govt. Social insurance	Private providers

<b>United Kingdom (1994/95)</b>					
Revenue collection	General taxation			Social insurance	Out-of-pocket
Pooling	Ministry of health				No pooling
Purchasing	Health authorities			GPs	Individual purchasing
Provision	National health service			Private providers	

Note: Widths are proportional to estimated flows of funds.  
Source: National health accounts estimates.

pooling arrangements in developing countries. Although an improvement over out-of-pocket financing, their size and organizational capacity often threatens their financial sustainability (16, 17). Predominantly out-of-pocket financing represents the highest degree of fragmentation. In such a case, each individual constitutes a pool and thus has to pay for his or her own health services.

Larger is better for pooling and purchasing. But economies of scale show diminishing returns and, beyond a critical size, marginal benefits may be negligible. The argument for large pools is therefore not an argument for single pools when multiple pools can exist without fragmentation, and when their size and financing mechanisms allow for adequate spreading of risk and subsidization of the poor.

Health system policy with regard to pooling needs to focus on creating conditions for the development of the largest possible pooling arrangements. Where a particular country for the moment lacks the organizational and institutional capacity to have a single pool or large pools for all citizens, policy-makers and donors should try to create the enabling conditions for such pools. Meanwhile, policy-makers should promote pooling arrangements whenever possible, as a transitional stage towards the future aggregation of pools. Even small pools or pools for segments of the population are better than pure out-of-pocket financing for all. Opposing or neglecting such arrangements until the capacity exists for the establishment of an effective single pool has two drawbacks. It deprives consumers of improved protection. And it may prevent the state from regulating such initiatives and steering them towards future large or single pool arrangements. Introducing regulations such as community rating (adjusting for the average risk of a group), portable employment-based pooling (insurance that a worker keeps when changing jobs) and equal minimum benefit packages (access to the same services in all pools), in addition to protecting members of the pools, may pave the way for larger pooling in the future.

For low income economies where the formal sector is small, this means promoting pooling at the community level. Communities' lack of trust in local pooling organizations might be a limiting factor, but such initiatives offer an important opportunity for international cooperation whereby donors act as guarantor for the community and help create the necessary organizational and institutional capacity. For middle income developing countries, this means both encouraging the creation of pools and, where possible, either directly establishing a large pool or enacting regulation to specify a minimum size of pool for financial viability, as well as regulating pooling initiatives in a way that will facilitate consolidation in the future.

However, competition among pools is not entirely bad. It can increase the responsiveness of pooling organizations to their members and provide an incentive for innovation. It can also offer incentives for reducing costs (to increase market share and profits), for example through mergers, as in the reform of the quasi-public health insurance organizations (*Obras Sociales*) in Argentina in 1996. Lack of competition meant that the administrators were little concerned about high administrative costs and small benefits for their members, as they had in any case a captive group of contributors. Competition and the resulting mergers, together with explicit subsidies for low-income beneficiaries, have allowed members of small pools to join larger pools and obtain better benefits for the same level of contributions.

Despite its potential benefits, pooling competition poses significant problems to health systems, particularly in selection behaviour by both pooling organizations and consumers. Mandatory participation (that is, all eligible members must join the pooling organization) significantly reduces the scope of selection behaviour but does not totally eliminate the

incentives associated with it, particularly under non-risk-related contribution schemes.

Selection behaviour is a potential problem of competition whenever and at whatever organizational level pooling is performed (27, 28). It is particularly a problem for competition under non-risk-related contribution schemes. Either pooling organizations will try to pick the lowest risk consumers (risk selection), who will contribute but not cause expense, or the highest risk consumers will seek coverage more actively than the rest of the population (adverse selection). Pooling competition then becomes a battle for information between consumers (who usually know more about their own risk of requiring health interventions) and the pooling organization (which needs to know more about consumers' risks to ensure long term financial sustainability). This has significant consequences for the administrative costs of pooling organizations. If adverse selection predominates, pooling organizations end up with increasing costs, are obliged to demand increasing contributions, and may eventually face financial default. This applies not only to private health insurance schemes but also to community pooling arrangements. Evidence shows that managing adverse selection is a major challenge for community pooling arrangements (17), which mostly rely on voluntary affiliation. If instead risk selection predominates, as is most likely when there is weak regulation of pooling competition, the poor and the sick will be excluded.

Exclusion from the pool is a problem that should be corrected through a combination of regulation and financial incentives. Regulation may cover such aspects as mandatory participation, non-risk-related contributions or community rating (the same price for a group of members sharing the same geographical area or the same workplace), and prohibition of underwriting (requesting additional information regarding health risks). Financial incentives may include risk compensation mechanisms and subsidies for the poor to join a pool. These approaches reduce the problems of pooling competition but are administratively expensive because of the high transaction costs within the system, associated with moving from hierarchical organizational arrangements for non-competitive pools to a market in pooling (29, 30).

Regulation and incentives should also be directed to avoiding fragmentation of the pool as a result of competition. If organizational and institutional incentives are adequate, large pools are much more efficient than pooling competition. Single national pools, as the largest pools attainable and as non-competing organizations, might be seen as the most efficient way to organize pooling. They avoid fragmentation and all competition problems but also forego the advantages of competition.

In most health financing arrangements, pooling and purchasing are integrated within the same organization. Allocation of funds from pooling to purchasing occurs in the organization through the budgetary process. There are, however, a few instances in the world where attempts have been made to separate the functions and allocate resources from a pooling organization to multiple purchasers through risk adjusted capitation. For example, in Colombia (31, 32) and the USA (33, 34), attempts have been made to take advantage of purchasing competition to minimize the pooling competition problems discussed above.

## STRATEGIC PURCHASING

Health systems need to ensure that the package of health interventions they provide and finance responds to the criteria discussed in Chapter 3. They also need to ensure that the way interventions are provided helps to improve the system's responsiveness and financial fairness. Strategic purchasing is the way to achieve this.

But, as shown in Chapter 3, the burden of ensuring the effectiveness of health interventions rests mainly on the shoulders of providers. To play their role effectively, providers need adequate inputs and organizational arrangements, as well as coherent incentives, both within and from outside the organization. Purchasing plays a central role in ensuring coherence of external incentives for providers through contracting, budgeting and payment mechanisms.

Strategic purchasing faces three fundamental challenges: What interventions to buy? From whom to buy them? And how to buy them? Size is also important for purchasing organizations. Large purchasers can not only take advantage of economies of scale but also of better bargaining capacity (monopsony power) regarding price, quality and opportunity of services, in dealing with natural monopolies on the provider side.

Strategic purchasing requires a continuous search for the best interventions to purchase, the best providers to purchase from, and the best payment mechanisms and contracting arrangements to pay for such interventions. Identifying the best providers means getting the best deals (for example, fast access for patients to the contracted services). It means establishing strategic alliances for the future development of those providers and for disseminating their best practices to other providers.

The important role of public health and the technical characteristics of what interventions to provide are discussed in Chapter 3. In purchasing personal care, the determination of what interventions to buy takes place at two levels. One level is largely related to stewardship. Here, society determines (most of the time implicitly) the relative weighting of the goals of the system – health, responsiveness, and fair contribution to financing. It does so by determining priorities for the public financing of specific programmes, or via regulation and financial incentives for voluntary or mandated private financing. In the presence of weak stewardship, the relative weighting of health system goals is defined *de facto* by the purchaser and the market forces. The second level is the purchaser's responsibility. This means that the purchaser is responsible for the day-to-day identification of the interventions to achieve the system goals (as defined at the stewardship level), as well as the determination of co-payment and other financial aspects. It also means that the purchaser has authority for negotiating with providers with regard to the expected quantity, quality, and availability of the interventions to be purchased and provided.

Purchaser organizations also need to define from whom to buy. This definition is crucial in allowing them to avoid becoming involved in the micro-management of providers. In order to set incentives for cost control, an emphasis on preventive care, and maintaining or improving the quality of services, purchasers need to prioritize among *units* of purchasing: that is, whether to buy individual interventions, specified packages of care, all the care for individuals or groups, or all the inputs needed for that care. Each unit of purchasing needs to be of a critical size, and to include a wide enough diversity of individual providers to ensure an appropriate mix of services. Such units make it easier for the purchaser and the provider to agree on a payment mechanism in which the provider shares the risk with the purchaser (that is, the provider is partly responsible for a full range of interventions for a relatively fixed amount of money). The spectrum of risk sharing, from all the risk borne by the purchaser to all of it transferred to providers, is discussed in *The world health report 1999*.

With such units, it is also easier for purchasers to make long-term contracts with providers who would take care of all aspects of necessary health care for groups of members of the pool. If the purchasing unit is too small, the purchaser will have difficulty in agreeing on a risk sharing payment mechanism, because of the potential fragmentation of the pool, and will have to resort to traditional input purchasing or fee-for-service. Such a situation will

force the purchaser to focus on short-term isolated interventions, as the absence of a risk sharing agreement will make it difficult to conclude a long-term contract for interventions for groups of the population. This will increase the overall administrative costs in the system relative to the volume of interventions involved.

With regard to how to buy, there are two objectives. The first is to avoid micro-purchasing, that is, such small scale buying of interventions that it constitutes the micro-management of providers. (There are, however, circumstances where micro-purchasing or micro-management may be justified, particularly for high complexity, very expensive and low frequency interventions.) The second is to design and implement effective contractual, budgeting and payment mechanisms. Avoiding micro-purchasing implies focusing the provisioning process on setting the right external incentives and evaluating results. The challenge here is to set purchasing goals that allow providers all necessary discretionary power in the provider–citizen contact, but which leave the purchaser the capacity to influence overall access to personal and non-personal services for members of the pool.

The budgeting and provider payment mechanisms are an essential part of the purchaser–provider interaction. Together with contracting, they establish an environment in which there are incentives for providers to act in accordance with the following four objectives: to prevent health problems of members of the pool; to provide services and solve health problems of members of the pool; to be responsive to people’s legitimate expectations; and to contain costs.

No single budgeting or provider payment mechanism can achieve all four objectives simultaneously (35). Table 5.3 summarizes the characteristics of the most common budgeting and payment mechanisms designed to meet those objectives. While line item budgets can be effective in controlling costs, they provide few incentives to achieve the other three objectives. In contrast, while fee-for-service provides strong incentives to deliver services, it also provides incentives that lead to an overall increase in the cost of the system. Therefore, purchasers need to use a combination of payment mechanisms to achieve their objectives. Free choice of provider by consumers increases responsiveness under all payment systems, but particularly under those needing to attract patients to ensure payment by the purchaser (fee-for-service or diagnostic related payment).

Capitation means a fixed payment per beneficiary to a provider responsible for delivering a range of services. It offers potentially strong incentives for prevention and cost control, to the extent that the provider receiving the capitation will benefit from both. If the contract is so short that a particular preventive intervention would have a noticeable effect only beyond the duration of the contract, there will be little or no incentive for prevention.

**Table 5.3 Provider payment mechanisms and provider behaviour**

<b>Provider behaviour</b>	Prevent health problems	Deliver services	Respond to legitimate expectations	Contain costs
<b>Mechanisms</b>				
Line item budget	+/-	--	+/-	+++
Global budget	++	--	+/-	+++
Capitation (with competition)	+++	--	++	+++
Diagnostic related payment	+/-	++	++	++
Fee-for-service	+/-	+++	+++	---

Key: +++ very positive effect; ++ some positive effect; +/- little or no variable effect; -- some negative effect; --- very negative effect.

Similarly, if the provider is not allowed to benefit from or reinvest the surplus resulting from savings, there is little incentive for cost control beyond that required for the financial sustainability of the provider organization.

Because of its advantages in cost control and prevention, capitation has been introduced in many purchasing organizations in the world. It has been used in the UK national health service with regard to general practitioners and later played a more important role in sharing risk with the introduction of general practitioner fundholding, allowing surpluses to be invested in the fundholder's practice (6). It has also been used for provider networks in Argentina's social security organization for retirees (23), in New Zealand with independent practice associations (36), and in the United States with health maintenance organizations (37). When risk-sharing payment mechanisms are used, depending on the specific terms of the payment mechanism, part of the pooling function of spreading risk among members of the pool may be performed by the provider. Thus, when an integrated pooling/purchasing organization contracts with smaller providers, each provider may also become a pooling organization. There is thus a risk of fragmenting the pool if the provider groups are too small. This has been the main argument for shifting from general practitioner fundholding to larger pools, the primary health care groups, in the UK in 1999.

Supply side provider payment mechanisms, such as line item budgets, focus purchasing efforts on inputs and make it impossible for providers to respond flexibly to external incentives. Too often these are the main resource allocation mechanisms for public providers in developing countries. As a result, providers do not continuously adapt their mix of services. This has been a serious barrier to improving health system efficiency in many developing countries (38). It has also been a major obstacle to the improvement of public-private collaboration in the provision of services (39). Line item budgets are in these respects much worse than global budgets, which also control costs.

What does moving to more flexible resource management at the provider level require? *The world health report 1999* introduced an answer to this question (14): it means reaching more explicit agreements between purchasers and providers regarding services to be provided (performance agreements, quasi-contracts and contracts). Quasi-contractual arrangements refer to non-legally-binding explicit agreements between two parties, in this case between the purchaser and the provider. Resource management also requires the introduction of "money follows the patient" schemes, particularly where policies favouring the free choice of providers are introduced. Doing it well demands significant organizational and institutional capacity, along with propitious political conditions, particularly because of the potential consequences for public providers. Failure to develop such capacity and political conditions before or simultaneously with entering into contracting and demand side financing reforms can have negative consequences to judge from experience in India, Mexico, Papua New Guinea, South Africa, Thailand and Zimbabwe (40, 41). Contracting out clinical services is particularly complex even when limited to non-profit providers such as church hospitals in Ghana, the United Republic of Tanzania and Zimbabwe (42).

In summary, purchasers need to move from supply side payment to demand side provider payment mechanisms, from implicit to explicit contracting, and from fee-for-service to some form of risk sharing payment mechanisms. Contracting, shifting to demand side payment, and introducing risk sharing provider payment mechanisms require a high level of technical, organizational and institutional capacity, as well as significant political leverage because of the likely resistance of providers to bearing more risk and being held more accountable, particularly in the public sector.

## ORGANIZATIONAL FORMS

The debate on policy alternatives for health system financing often focuses exclusively on technical aspects, underestimating the importance of organizational and institutional factors. Examples of the results of this approach include the provider payment mechanism reforms designed in the early 1990s in some Latin American countries (Argentina, Chile, Costa Rica, and Nicaragua) (39). These reforms initially underestimated the importance of organizational and institutional effects, assuming that having the right price signals would be sufficient to change provider behaviour. It seems to have been assumed (explicitly or implicitly) that managers of public providers would – mainly by virtue of such new mechanisms as diagnostic-related payments or capitation – understand the price signals, know how to respond and be willing to act accordingly, despite the culture of their organizations. These reforms also underestimated the importance of and difficulties involved in providing managers with a flexible enough legal and administrative environment to make the correct changes. Furthermore, the reforms seem to have assumed that the government would be willing and able to deal with the political problems associated with such flexibility. Experience over the last 10 years shows that these assumptions are not always correct, and that more emphasis on organizational and institutional change is required to make provider payment reforms work.

Characteristics of provider organizations are analysed in Chapter 3. A similar analysis is valid for health financing organizations. Some of the most important factors affecting the performance of health financing organizations and, through it, the financial risk protection provided by the health system are discussed below.

In addition to contributing to the health system via out-of-pocket payment at the moment of demanding services, citizens also contribute to most health systems in the world through various combinations of the following *organizational forms*.

- *Ministry of health*, usually heading a large network of public providers organized as a national health service, relying on general taxation – collected by the ministry of finance – as the main source of revenue, and serving the general population.
- *Social security organization* (single or multiple, competing or not), mostly relying on salary-related contributions, owning provider networks or purchasing from external providers, and serving mostly their own members (usually formal sector workers).
- *Community or provider based pooling organization*, usually comprising a small pooling/purchasing organization relying mostly on voluntary participation.
- *Private health insurance fund* (regulated or unregulated), mostly relying on voluntary contributions (premiums), which may be risk-related but are usually not income-related, and are often contracted by an employer for all a firm's employees.

Providers can play a role as pooling organizations under a non-risk-adjusted capitation payment mechanism, as discussed above. In this scenario, internal incentives for providers coexist with internal incentives for financing organizations, which may impede coherence among incentives.

Each organizational form deals with the technical characteristics of health financing in a particular way. This is particularly evident in comparing private risk-related health insurance with social security. Social security organizations spread risk among the whole pool through non-risk-related contributions. All members of the pool pay a proportion of their salary, regardless of their risk. In contrast, voluntary private health insurance contributions charge the same premium only for the members of a similar risk category in the pool (such

as the same sex, age and place of residence). There are multiple categories in private health insurance, and members are charged according to the risk category to which they belong. The social security and risk-related private insurance approaches are contradictory, and their coexistence creates different incentives for consumers. All consumers whose risk category is such that private insurance would charge them less than the amount that they would have to pay under social insurance have the incentive to avoid contributing to social insurance and use private insurance if they are allowed to. High-risk people, however, have the incentive to contribute to social security, loading it with high-risk members and increasing the per capita cost of services for members of the pool. The Chilean case, presented in Box 5.2, is an example of this phenomenon (43), in which contributors can opt out of social security and direct their contributions to private insurers. The contradictory incentives can be controlled only if social insurance is mandatory.

Health financing functions are often integrated in a single organization. For ministries of health (or national health services), however, collecting is usually done by the ministry of finance. Some health systems with multiple social security organizations have introduced central collecting agencies in charge of risk equalization among pools (as in Colombia, Germany, and the Netherlands). Various attempts have been made to separate the pooling

### Box 5.2 The Chilean health insurance market: when stewardship fails to compensate for pooling competition problems and for imbalances between internal and external incentives

In 1980, Chile implemented a radical reform of the health system. It separated financial administration in the public health sector from public providers and the Ministry of Health, creating the National Health Fund (FONASA), which is financed by a combination of general taxation (for the poor who also are included in the pool) and a 7% payroll tax contribution for formal sector workers. It simultaneously allowed for the introduction of private competing health insurance organizations (ISAPREs). All formal sector workers and their families have to contribute either to FONASA or an ISAPRE. All the rest of the population is covered by FONASA. In contrast with FONASA which charges all members the same 7% payroll tax irrespective of the risk, ISAPREs are allowed to adjust the contribution (with the 7% payroll tax as a minimum contribution) and the benefit package to the risk of the principal and his or her family. These organizational forms reflect opposing rationales. While FONASA is based on salary-related contributions with no exclusions, ISAPREs in practice are based on

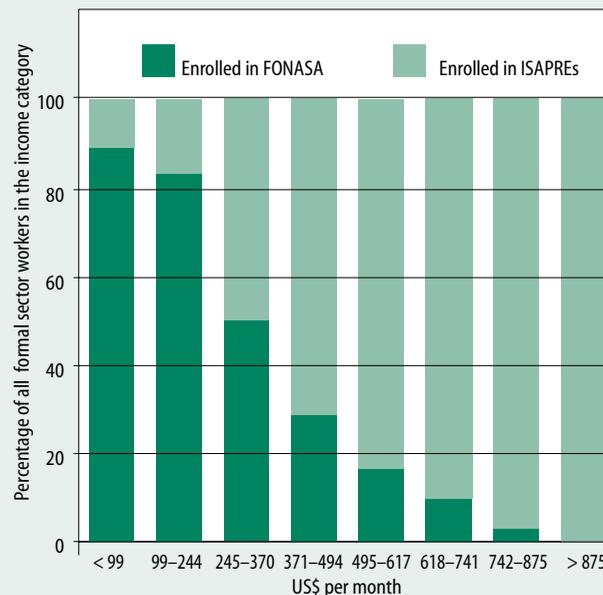
risk-related contributions. Apart from the very limited power of the Ministry of Health, no regulatory agency was in a position to regulate ISAPREs until 10 years after they were created. As a result, ISAPREs grew from covering 2% of the popu-

lation in 1983 to 27% in 1996.

Lack of regulation, weak stewardship (for political reasons), and an explicit policy to channel all cross-subsidies through FONASA only, resulted in severe segmentation of the market. ISAPREs focused on the

richest, and risk-selected the healthiest. Only recently has it been possible to introduce regulation to reduce risk selection. Segmentation has determined that while more than 9% of the total Chilean population is older than 60 years of age (generally the highest risk group in the population), that population group represents only about 3% of ISAPRE beneficiaries. At the same time, as shown in the graph, while almost all low income workers are in FONASA, very few are in the ISAPRE system. There is continued debate in Chile over reform of the health insurance system to address this structural problem.

#### Health insurance of formal sector workers, enrolment in FONASA and ISAPREs by income level, Chile, 1994



Source: ISAPREs Association and FONASA, 1995.

Source: Baeza C, Copetta C. *Análisis conceptual de la necesidad y factibilidad de introducir mecanismos de ajuste de riesgo y portabilidad de los subsidios públicos en el sistema de seguros de salud en Chile.* [Conceptual analysis of the necessity and feasibility of introducing mechanisms for risk adjustment and portability of public subsidies in the health insurance system of Chile.] Santiago, Chile, Centro Latinoamericano de Investigación para Sistemas de Salud (CLAISS) and Fondo de Promoción de Políticas Públicas de la Universidad de Chile, 1999 (in Spanish).

and the purchasing functions (as in Colombia and the United States). The organizational separation of collecting and pooling is less frequent than the separation between purchasing and provision, and it has been less explored. It appears to be less important for setting the right incentives for providers than the separation between purchasing and provision as introduced under managed competition and internal market reforms (44–47).

## INCENTIVES

As for provider organizations described in Chapter 3, health financing organizations are subject to internal incentives. Organizational performance depends on the coherence of the following internal incentives.

- The level of *autonomy* or decision rights that the organization has *vis-à-vis* its owner, its overseeing authority or the government. Critical decision rights include setting contribution levels (premiums or payroll tax), co-payment levels, prioritization of interventions to be purchased, designing and negotiating contracts and provider payment mechanisms, selectivity in contracting with providers, and in many cases, freedom to determine investments.
- The degree of *accountability*. As autonomy increases, owners, overseeing authorities or the government require mechanisms to make the organization responsible for the expected results via hierarchical supervision, regulation or financial incentives.
- The degree of *market exposure*, that is, the proportion of revenues earned in a competitive way rather than acquired through a budget allocation. Particularly important for performance is whether governments provide budget supplements for deficits that originate from poor performance.
- The degree of *financial responsibility* for losses, and rights to profits (retained earnings and proceeds from the sale or rental of capital).
- The degree of *unfunded mandates*, that is, the proportion (in terms of revenues allocated) of mandates for which the organization is legally held responsible but for which it is not allowed to charge fees, and for which the organization does not receive any compensatory financial transfer. Such mandates may be to include the very poor or the very sick in the pool, as is usually the case for ministries of health or national health services. There may also be a mandate for the purchaser to pay for emergency care in a life-threatening situation, no matter where the care is provided and whatever the cost.

All prepaid health financing systems in the world are composed of combinations of the four organizational forms described above. It is clear that each organizational form has a different level of exposure to internal incentives. For example, ministry of health or ministry of finance organizations are much more likely to bear unfunded mandates than private insurance funds. Furthermore, because of the differences in market exposure and accountability between such organizations, their responses to unfunded mandates will be significantly different. While ministries of health or finance can respond to unfunded mandates by adjusting the quality or opportunity of interventions or even generating budget deficits, private insurance funds might respond by excluding members who are at a high risk of requiring the services required by the unfunded mandates. To avoid negative equity consequences, particularly under increasing autonomy, regulatory and financial incentives (e.g. risk compensation mechanisms) are necessary to protect the sick and the poor.

Another example of the significant differences in internal incentives concerns to whom each organizational form is accountable. Because ministries of health or finance are accountable to government, external incentives are required to make sure that they are also responsive to consumers. On the other hand, because private health insurance is accountable to owners and consumers, external incentives and regulation are needed to make sure that benefit packages and insurance practices are coherent with national priorities and policies regarding health, financial fairness and responsiveness. Often, as was the case with unregulated private health insurance in Argentina until 1996 (23), private insurance responds to consumer demand by focusing benefit packages on low-cost and high frequency interventions, and excluding very high-cost and low frequency interventions (catastrophic events) which are most appropriately included in pooling arrangements. Regulating minimum benefits for all members, including coverage of catastrophic events by each fund or through re-insurance, is necessary in these circumstances.

Table 5.4 summarizes the level of each internal incentive for each of the four organizational forms.

To increase health system performance, stewardship has a major role to play in health financing. This is because external incentives are needed to compensate for differences in the internal incentives faced by the different health financing organizations.

A set of external incentives (rules and customs) governs the way the different organizational forms interact within the system. The three key external incentives that influence the behaviour of health financing organizations are the rules and customs relating to governance, public policy objectives, and control mechanisms.

- The rules and customs relating to *governance* shape the relationship between organizations and their owners. Ownership (public or private) usually provides the right to make decisions over the use of an asset and the right to the income that remains after all fixed obligations are met. Specification and limitation of these rights is often a major element of regulation.
- The rules and customs related to *public policy objectives* that influence the behaviour of organizations include budget implementation directives (for ministries of health or national health services), criteria for eligibility for public subsidies (for private insurers and community pools), and required auditing procedures.
- The rules and customs relating to *control mechanisms* shape the relationships between organizations and the public authorities, as well as between organizations

**Table 5.4 Exposure of different organizational forms to internal incentives**

<b>Organizational forms</b> <b>Internal incentives</b>	Ministries of health or finance	Social security organizations	Community pooling organizations	Private health insurance funds
Decision rights (autonomy)	Limited	Variable but usually high	High	High
Accountability	Government, voters	Board/often government	Owners / consumers	Owners / consumers
Market exposure	None	Variable; high when multiple organizations compete	High	High
Financial responsibility	None or very limited	Low	High	High
Unfunded mandates	High	Low	None or very limited	None or very limited

and consumers. In this context, the public authorities are those involved in areas such as policy-making, regulation and enforcement. The public authorities have a range of instruments at their disposal with which to set external incentives for health financing organizations, ranging from hierarchical command and control (e.g. political or administrative instructions from the government to the ministry of health or national health service) to regulation and financial incentives. These instruments may include rules related to such subjects as the percentage of payroll tax to be devoted to financing social security organizations, the minimum contents of benefit packages, allowed exclusions and pre-existing conditions which must be covered, duration of contracts, commercialization and marketing restrictions, the pricing of private insurance, and the mandatory sending of information to the regulatory agencies.

As for internal incentives, the four organizational forms are subject to different degrees of exposure to the various external incentives. Table 5.5 summarizes the most important differences.

The difference between the external incentives for ministries of health or finance and private health insurance funds is particularly relevant. While hierarchical control influences ministries of health or finance, it has little or no influence on private insurance or community pooling arrangements. The introduction of private competitive health insurance (as an explicit policy option) or the growth of informal community pooling arrangements (or informal health insurance) require stewardship to shift from hierarchical control to using regulations and financial incentives as a means of influencing behaviour. This shift usually represents a significant change in the way control has traditionally worked. It requires an ability to anticipate and implement the necessary legal and administrative changes, and it demands a significant alteration in the skill mix and culture of control organizations.

Evidence from trends in health financing reforms in some eastern European and Latin American countries (3, 48) shows the potential negative effects of failure to strengthen control and shift to different external incentive instruments when private competitive health insurance is introduced. Risk selection is almost certain, taking high income low-risk consumers out of the public pools and worsening the financial situation of the latter.

To realize their potential, external and internal incentives should be coherent and aligned to address two fundamental problems increasingly evident in developing countries: the decision-making process being “captured” by other interests; and inefficiencies in supply side financing.

**Table 5.5 Exposure of different organizational forms to external incentives**

<b>Organizational forms</b> <b>External incentives</b>	Ministries of health or finance	Social security organizations	Community pooling organizations	Private health insurance funds
Governance	Public, low level of decision rights	Public or quasi-public with variable levels of decision rights	Private, high level of decision rights	Private, high level of decision rights
Financing for public policy objectives	High	Variable; government and market	None, except when receiving conditional public subsidies	None, except when receiving conditional public subsidies
Control mechanisms	Hierarchical control	Variable degrees of hierarchical control, regulations and financial incentives	Regulations and possibly financial incentives	Regulations and possibly financial incentives

As internal and external incentives make ministries of health or finance and even single social security organizations focus more on political concerns than on the interests of consumers, these organizations are particularly vulnerable to capture. In other words, decision-making in the pooling or purchasing organization is driven by interests other than health, responsiveness to beneficiaries and financial fairness. Capture may happen as a result of fiscal interests, corporate interests, union interests, political party interests, and so on. There are many examples of systems where social security revenues are used for fiscal purposes (a common problem in Latin America in the past) or where the government, as an employer, simply does not pay its social security dues under tripartite financing arrangements (workers, employers and government all contribute), as in Costa Rica during the 1980s. Strikes by physicians and their effects on salaries in national health services also show the vulnerability of such systems to capture by professional interests and illustrate one danger of large-scale public provision.

## HOW FINANCING AFFECTS EQUITY AND EFFICIENCY

The most important determinant of how fairly a health system is financed, as illustrated in Chapter 2, is the share of prepayment in total spending. Out-of-pocket payment is usually the most regressive way to pay for health, and the way that most exposes people to catastrophic financial risk. How revenues are collected therefore has a great impact on the equity of the system.

But even if nearly any form of prepayment is preferable, on these grounds, to out-of-pocket spending, it also matters greatly how the revenues are combined so as to share risks: how many pools there are, how large they are, whether inclusion is voluntary or mandatory, whether exclusion is allowed, what degree and kind of competition exists among pools, and whether, in the case of competing pools, there are mechanisms to compensate for differences in risk and in capacity to pay. All these features affect the fairness of the system, but they also help determine how efficiently it operates. The argument in favour of a single pool or a small number of pools of adequate size, and against fragmentation, concerns the financial viability of pools, the administrative costs of insurance, the balance between the economies of scale and (when there is little or no competition) the risks of capture and unresponsiveness, and the limitation of risk selection (which is a matter of efficiency as well as equity). Inefficiencies in collecting and pooling revenues reduce both the funds available for investment and for providing services, and people's access to those services that can be financed.

Purchasing, finally, also affects both equity and efficiency, by determining which investments are made and which interventions are bought, and for whom. Revenues may be collected fairly and with minimal waste, and be pooled so as to assure that the healthy help support the sick and the rich help support the poor. The performance of the system will still fall short of its potential if the pooled resources are not used intelligently to purchase the best attainable mixture of actions to improve health and satisfy people's expectations.

## REFERENCES

1. **Musgrove P.** Reformas al sector salud en Chile: contexto, logica y posibles caminos [Reforms to the Chilean health sector: context, logic and possible trajectories]. In: Giaconi J, ed. *La salud en el siglo XXI*. Santiago, Chile, Centuro de Estudios Publicos, 1995 (in Spanish).
2. **Oyarzo C, Galleguillos S.** Reforma del Sistema de Salud Chileno: Marco Conceptual de la Propuesta del Fondo Nacional de Salud [Reform of the Chilean health system: conceptual framework of the National Health Fund proposal]. *Cuadernos de Economia*, 1995, **32**(95) (in Spanish).
3. **Preker A et al.** *Health financing systems in transition: trends in Eastern Europe and Central Asia*. Washington, DC, The World Bank, 1999 (World Bank Technical Paper, forthcoming).
4. **New Zealand Ministry of Health.** *Healthy New Zealanders: briefing papers for the Minister of Health 1996*. Wellington, Ministry of Health, 1996.
5. **Hornblow A.** New Zealand's health reforms: a clash of cultures. *British Medical Journal*, 1997, **314**: 1892–1894.
6. **Robinson R, Le Grand J.** *Evaluating the National Health Service reforms*. Oxford, Policy Journals, Transaction Books, 1994.
7. **UK Department of Health.** *Working for patients*. London, HMSO, 1989.
8. **UK Department of Health.** *The new NHS*. London, The Stationery Office, 1998.
9. **Gertler P, van der Gaag J.** *The willingness to pay for medical care: evidence from two developing countries*. Baltimore, MD, The Johns Hopkins University Press, 1990.
10. **Bitran R, McInnes DK.** *The demand for health care in Latin America: lessons from the Dominican Republic and El Salvador*. Washington, DC, Economic Development Institute of The World Bank, 1993.
11. **Lavy V, Quingley JM.** *Willingness to pay for the quality and intensity of medical care by low-income households in Ghana*. Washington, DC, The World Bank, 1993 (Living Standards Measurement Study, Working Paper No. 94).
12. **Nyonator F, Kutzin J.** Health for some? The effects of user fees in the Volta Region of Ghana. *Health Policy and Planning*, 1999, **14**(4): 329–341.
13. **Musgrove P.** *Public and private roles in health: theory and financing patterns*. Washington, DC, The World Bank, 1996 (World Bank Discussion Paper No. 339).
14. *The world health report 1999 – Making a difference*. Geneva, World Health Organization, 1999: 37–43.
15. **Shieber G, Maeda A.** A curmudgeon's guide to financing health care in developing countries. In: *Innovations in health care financing*. Proceedings of a World Bank Conference, Washington, DC, 1997.
16. **Bennett S, Creese A, Monash R.** *Health insurance schemes for people outside formal sector employment*. Geneva, World Health Organization, 1998 (document WHO/ARA/CC/98.1).
17. **ILO.** *Estudio de casos de esquemas de extension de cobertura en salud para el sector informal en America Latina [Case studies of schemes for extending health coverage for the informal sector in Latin America]*. Santiago, Chile, Centro Latinoamericano de Investigación para Sistemas de Salud (CLAISS), 1999 (in Spanish).
18. **Newhouse JP, Manning WG, Morris CN.** Some interim results from a controlled trial of cost sharing in health insurance. *New England Journal of Medicine*, 1981, **305**: 1501–1507.
19. **Abel-Smith B, Rawal P.** Can the poor afford 'free' health services? A case study of Tanzania. *Health Policy and Planning*, 1992, **7**(4): 329–341.
20. **Schefler R, Yu W.** Medical savings accounts: a worthy experiment. *European Journal of Public Health*, 1998, **8**: 274–276.
21. **Saltman RB.** Medical savings accounts: a notably uninteresting policy idea. *European Journal of Public Health*, 1998, **8**: 276–278.
22. **Londoño JL, Frenk J.** Structured pluralism: towards an innovative model for health system reform in Latin America. *Health Policy*, 1997, **41**(1): 1–36.
23. **World Bank.** *Argentina, facing the challenge of health insurance reform*. Washington, DC, The World Bank, 1997 (LASHD ESW Report 16402-AR).
24. **Baeza C, Cabezas M.** *Es necesario el ajuste de riesgo en los mercados de seguros de salud en América Latina? [Is risk adjustment necessary in health insurance markets in Latin America?]*. Santiago, Chile, Centro Latinoamericano de Investigación para Sistemas de Salud (CLAISS), 1999 (in Spanish).
25. **Van de Ven WP et al.** Risk-adjusted capitation: recent experiences in The Netherlands. *Health Affairs*, 1994, **13**(5): 120–136.

26. **Musgrove P.** *Equitable allocation of ceilings on public investment: a general formula and a Brazilian example in the health sector.* Washington, DC, The World Bank, 1996 (Working Paper No. 69).
27. **Arrow K.** Uncertainty and the welfare economics of medical care. *American Economic Review*, 1963, 53(5): 941–973.
28. **Feldstein P.** *Health care economics.* New York, Delmar Publisher, 1993.
29. **Coase R.** The nature of the firm. *Economica*, 1937, 4.
30. **Williamson O.** *The economic institutions of capitalism.* New York, The Free Press, 1985.
31. **Londoño JL.** Estructurando pluralismo en los servicios de salud: la experiencia Colombiana [Structuring pluralism in health services: the Colombian experience]. *Revista de Análisis Económico*, 1996, 11(2) (in Spanish).
32. **Titelman D, Uthoff A.** *The health care market and reform of health system financing.* Santiago, Chile, ECLA, 1999.
33. **Weiner J et al.** Risk-adjusted Medicare capitation rates using ambulatory and inpatient diagnoses. *Health Care Financing Review*, 1996, 17(3): 77–99.
34. **Actuarial Research Corp.** *Medical capitation rate development.* Annandale, VA, US Department of Commerce, 1996 (National Technical Information, 1996 PB 96214887).
35. **Barnum H, Kutzin J, Saxenian H.** Incentives and provider payment methods. *International Journal of Health Planning and Management*, 1995, 10: 23–45.
36. **Wilton P, Smith RD.** Primary care reform: a three country comparison of budget holding. *Health Policy*, 1998, 44(2): 149–166.
37. **Sekhri Feachem N.** Managed care: the US experience. *Bulletin of the World Health Organization*, 2000, 78(6) (in press).
38. **Preker A, Feachem RGA.** *Market mechanisms and the health sector in Central and Eastern Europe.* Washington, DC, The World Bank, 1995 (Technical Paper No. 293).
39. **Baeza C.** *Taking stock of health sector reform in LAC.* Washington, DC, The World Bank, 1998 (World Bank Technical Discussion Paper, LASHD).
40. **Mills A.** Contractual relationships between governments and the commercial private sector in developing countries. In: Bennett S, McPake B, Mills A, eds. *Private health providers in developing countries: serving the public interest?* London, Zed Books, 1996.
41. **McPake B, Hongoro C.** Contracting out clinical services in Zimbabwe. *Social Science and Medicine*, 1995, 41(1): 13–24.
42. **Gilson L et al.** Should African governments contract out clinical health services to church providers? In: Bennett S, McPake B, Mills A, eds. *Private health providers in developing countries: serving the public interest?* London, Zed Books, 1996.
43. **Baeza C, Copetta C.** *Análisis conceptual de la necesidad y factibilidad de introducir mecanismos de ajuste de riesgo y portabilidad de los subsidios públicos en el sistema de seguros de salud en Chile* [Conceptual analysis of the necessity and feasibility of introducing mechanisms for risk adjustment and portability of public subsidies in the health insurance system of Chile]. Santiago, Chile, Centro Latinoamericano de Investigación para Sistemas de Salud (CLAISS) and Fondo de Promoción de Políticas Públicas de la Universidad de Chile, 1999 (in Spanish).
44. **Enthoven A.** *Reflections on the management of the National Health Service.* London, Nuffield Provincial Hospitals Trust, 1985.
45. **Enthoven A.** Managed competition of alternative delivery systems. *Journal of Health Politics, Policy and Law*, 1988, 13: 305–321.
46. **Enthoven A.** The history and principles of managed competition. *Health Affairs*, 1993, 12: 24–48.
47. **Ovretveit J.** *Purchasing for health: a multidisciplinary introduction to the theory and practice of health purchasing.* Buckingham, UK, Open University Press, 1995 (Health Services Management Series).
48. **Baeza C.** *Problemas y desafíos para el sistema de salud chileno en el siglo XXI* [Problems and challenges for the Chilean health system in the 21st century]. Santiago, Chile, Centro Latinoamericano de Investigación para Sistemas de Salud (CLAISS), 1999 (Documentos para el Dialogo en Salud No. 3) (in Spanish).

## CHAPTER SIX

# *How is the Public Interest Protected?*

*Governments should be the “stewards” of their national resources, maintaining and improving them for the benefit of their populations. In health, this means being ultimately responsible for the careful management of their citizens’ well-being. Stewardship in health is the very essence of good government. For every country it means establishing the best and fairest health system possible. The health of the people must always be a national priority: government responsibility for it is continuous and permanent. Ministries of health must take on a large part of the stewardship of health systems.*

*Health policy and strategies need to cover the private provision of services and private financing, as well as state funding and activities. Only in this way can health systems as a whole be oriented towards achieving goals that are in the public interest. Stewardship encompasses the tasks of defining the vision and direction of health policy, exerting influence through regulation and advocacy, and collecting and using information. At the international level, stewardship means influencing global research and production to meet health goals. It also means providing an evidence base to guide countries’ efforts to improve the performance of their health systems.*

## 6

## HOW IS THE PUBLIC INTEREST PROTECTED?

### GOVERNMENTS AS STEWARDS OF HEALTH RESOURCES

Stewardship is the last of the four health systems functions examined in this report, and it is arguably the most important. It ranks above and differs from the others – service delivery, input production, and financing – for one outstanding reason: the ultimate responsibility for the overall performance of a country’s health system must always lie with government. Stewardship not only influences the other functions, it makes possible the attainment of each health system goal: improving health, responding to the legitimate expectations of the population, and fairness of contribution. The government must ensure that stewardship percolates through all levels of the health system in order to maximize that attainment.

Stewardship has recently been defined as a “function of a government responsible for the welfare of the population, and concerned about the trust and legitimacy with which its activities are viewed by the citizenry” (1). It requires vision, intelligence and influence, primarily by the health ministry, which must oversee and guide the working and development of the nation’s health actions on the government’s behalf. Much of this chapter, therefore, addresses the ministry’s role.

Some aspects of stewardship in health must be assumed by government as a whole. Affecting the behaviour of health actors in other sectors of the economy, or ensuring the right size and skill mix of the human resources produced for the health system, may be beyond the ministry’s reach. The government ought to ensure coherence and consistency across departments and sectors, where necessary by an overall reform of public administration.

Outside of government, stewardship is also a responsibility for purchasers and providers of health services who must ensure that as much health as possible results from their spending. And stewardship in health has an international dimension, relating to external assistance.

But government remains the prime mover. Today in most countries the role of the state in relation to health is changing. People’s expectations of health systems are greater than ever before, yet limits exist on what governments can finance and on what services they can deliver. Governments cannot stand still in the face of rising demands. They face complex dilemmas in deciding in which direction to move: they cannot do everything. But in terms of effective stewardship, their key role is one of oversight and trusteeship – to follow the advice of “row less and steer more” (2, 3).

Stewardship has major shortcomings everywhere. This chapter examines some of them, then discusses important stewardship tasks. It considers the main protagonists involved, and strategies for implementing stewardship in different national settings. Finally, it brings together some of the messages from preceding chapters on policy directions for better-functioning health systems.

## WHAT IS WRONG WITH STEWARDSHIP TODAY?

“Ministries of health in low and middle income countries have a reputation for being among the most bureaucratic and least effectively managed institutions in the public sector. Designed and initiated in the early 20th century and given wide responsibility for financing and operating extensive public hospital and primary care systems in the post-war period, they became large centralized and hierarchical public bureaucracies, with cumbersome and detailed administrative rules and a permanent staff with secure civil service protections. The ministries were fragmented by many vertical programmes which were often run as virtual fiefdoms, dependent on uncertain international donor funding” (4).

The problems described above are familiar, in greater or lesser degree, in many countries today. The consequences are easy to see, but it is not always easy to see why the problems occur or how to solve them. Often that is because the stewards of health suffer specific visual impairments.

*Health ministries often suffer from myopia.* Because they are seriously short-sighted, ministries sometimes lose sight of their most important target: the population at large. Patients and consumers may only come into view when rising public dissatisfaction forces them to the ministry’s attention. In addition, myopic ministries recognize only the closest actors in the health field, but not necessarily the most important ones, who may be in the middle or far distance.

Ministries deal extensively with a multitude of public sector individuals and organizations providing health services, many of which may be directly funded by the ministry itself. Often, this involvement means intensive professional supervision and guidance. But sometimes just beyond their field of vision lie at least two other groups with a major role to play in the health system: nongovernmental providers, and health actors in sectors other than health.

In their size and potential impact on achieving health goals, these little recognized individuals and organizations may be more important than the public resources directed through the health ministry. Yet information about them may be scant, and a policy approach towards them is often lacking. In Myanmar, Nigeria (5), or Viet Nam, for example, privately financed and provided medical care is three or four times as big, in expenditure terms, as spending on public services. But the many different types of private providers in these countries are barely recognized in legislation and regulation.

Some large health insurance schemes in India currently have no legal status (6). In Europe and the Americas, road traffic accidents rank fourth in the total burden of disease. Yet the main involvement of the health ministry is often as a steward of accident and emergency services, not as a force for prevention. Services funded from public sources are obviously the responsibility of government. But private finance and the provision of *all* health actions clearly need to be within the focus of government as overall steward of the public interest.

Ministries are also myopic in the sense that their vision does not extend far enough into the future. Investment decisions – new buildings, equipment and vehicles – frequently

occupy the foreground, while the severe and chronic need to improve the balance between investment and recurrent funding fades into the hazy distance.

*Tunnel vision in stewardship* takes the form of an exclusive focus on legislation and the issuing of regulations, decrees, and public orders as means of health policy. Explicit, written rules have an important role to play in the performance of the stewardship function. But formulating regulations is relatively easy and inexpensive. It is also often ineffective, with ministries lacking the capacity to monitor compliance: there are seldom enough public health inspectors to visit all food shops and eating places or enough occupational safety inspectors to visit all factories regularly. On the rare occasions when sanctions are invoked they are too mild to discourage illegal practices or to affect widespread disregard of regulations.

Good stewardship needs the support of *several* strategies to influence the behaviour of the different stakeholders in the health system. Among these are a better information base, the ability to build coalitions of support from different groups, and the ability to set incentives, either directly or in organizational design. As authority becomes devolved, delegated and decentralized to a wide range of stakeholders in the health system, the repertoire of stewardship strategies needs to move away from dependence on “command and control” systems towards ensuring a cohesive framework of incentives.

*Health ministries sometimes turn a blind eye* to the evasion of regulations which they themselves have created or are supposed to implement in the public interest. A widespread example is the condoning of illicit fee collecting by public employees, euphemistically known as “informal charging”. A recent study in Bangladesh found that unofficial fee payments were 12 times greater than official payment (7). Paying bribes for treatment in Poland is cited as a common infringement of patients’ rights (8). Though such corruption materially benefits a number of health workers, it deters poor people from using services they need, making health financing more unfair, and it distorts overall health priorities.

In turning a blind eye, stewardship is subverted; trusteeship is abandoned and institutional corruption sets in. A blind eye is often turned when the public interest is threatened in other ways. For instance, doctors can remain silent through misplaced professional loyalty in the face of incompetent and unsafe medical practice by colleagues. A 1999 US study commented “whether care is preventive, acute or chronic, it frequently does not meet pro-

### Box 6.1 Trends in national health policy: from plans to frameworks

National health policy documents have a long history, predating but stimulated by international concern for promoting primary health care. In many centrally planned and developing economies, health policies were part of a national development plan, with a focus on investment needs. Some health policy documents were only a collection of project or programme-specific plans. They ignored the private sector and often took inadequate account of fi-

nancial realities and people’s preferences. Implementation problems were common.

By no means all countries have formal national health policies: France, Switzerland, and the United States do not; Tunisia has no formal single national policy document; the UK produced its first formal document in the 1990s; Portugal in 1998. The lifespan of a policy depends on whether there are fundamental changes to the agenda: India is still using its 1983 plan; Mongolia, in

economic transition, revised its 1991 policy in 1996 and again in 1998.

A shift is now occurring towards more inclusive – but less detailed – policy frameworks mapping the direction but not spelling out the operational detail, as in Ghana and Kenya.

*A national health policy framework:*<sup>1</sup>

- identifies objectives and addresses major policy issues;
- defines respective roles of the public and private sectors in fi-

ancing and provision;

- identifies policy instruments and organizational arrangements required in both the public and private sectors to meet system objectives;
- sets the agenda for capacity building and organizational development;
- provides guidance for prioritizing expenditure, thus linking analysis of problems to decisions about resource allocation.

<sup>1</sup> Cassels A. *A guide to sector-wide approaches for health development*. Geneva, World Health Organization/DANIDA/DFID/European Commission, 1997 (unpublished document WHO/ARA/97.12).

essional standards” (9). Ensuring probity in decisions on capital projects and other large purchasing decisions (equipment, pharmaceutical orders), where corruption may be particularly lucrative, is another frequent challenge to good stewardship.

Some recent developments create opportunities for better vision and more innovative stewardship. Greater autonomy in decisions relating to purchasing and service provision, for example, shifts some responsibility away from central or local government. But it creates new tasks for government in overseeing that both purchasing and provision are carried out in accordance with overall policy. Accumulated experience of practices such as contracting is now available (10) and rapid technological advances enable the fast, inexpensive handling of huge amounts of information, thus making it easier in principle for stewards to visualize the whole health system.

The notion of stewardship over all health actors and actions deserves renewed emphasis. Much conceptual and practical discussion is needed to improve the definition and measurement of how well stewardship is actually implemented in different settings. But several basic tasks can already be identified:

- formulating health policy – defining the vision and direction;
- exerting influence – approaches to regulation;
- collecting and using intelligence.

These tasks are discussed below.

## HEALTH POLICY – VISION FOR THE FUTURE

An explicit health policy achieves several things: it defines a vision for the future which in turn helps establish benchmarks for the short and medium term. It outlines priorities and the expected roles of different groups. It builds consensus and informs people, and in doing so fulfils an important role of governance. The tasks of formulating and implementing health policy clearly fall to the health ministry.

Some countries appear to have issued no national health policy statement in the last decade; in others, policy exists in the form of documents which gather dust and are never translated into action. Too often, health policy and strategic planning have envisaged unre-

### Box 6.2 Ghana’s medium-term health policy framework

In Ghana, after an extensive process of consultation, the following strategies were identified as providing the means to better performance in health.

- Re-prioritization of health services to ensure that primary health care services (i.e. services with maximum benefits in terms of morbidity and mortality reduction) receive more emphasis in resource allocation.
- The strengthening and decentralization of management within the context of a national health service.
- Forging linkages between private and public providers of health care to ensure consensus and that all resources are focused on a common strategy.
- Expansion and rehabilitation of health infrastructure to increase coverage and improve quality.
- Strengthening human resource planning, management and training as a means of providing and retaining adequate numbers of good quality and well-motivated health teams to provide the services.
- Provision and management of adequate logistics such as drugs and other consumables, equipment, and vehicles at all levels of the health system.
- Strengthening the monitoring and regulatory systems within the health service to ensure more effective implementation of programmes.
- Empowering households and communities to take more responsibility for their health.
- Improving the financing of health care by ensuring the efficient and effective use of all available resources from government, nongovernmental organizations, and private, mission and donor sources. Ways of mobilizing additional resources with a view to making the services more accessible and affordable will also be explored.
- Promoting intersectoral action for health development, particularly in the areas of food and nutrition, employment, education, water and sanitation.

Source: *Medium-term health strategy: towards vision 2020 Republic of Ghana*. Accra, Ministry of Health, 1995.

alistic expansion of the publicly funded health care system, sometimes well in excess of national economic growth. Eventually, the policy and planning document is seen as infeasible and is ignored. Box 6.1 sketches how comprehensive health planning has given way to a more flexible 'framework' approach. Ghana's 1995 medium-term health strategy identified ten ways in which the health system would contribute towards better health (see Box 6.2).

Public consultation occurs in some countries at the beginning of the policy formulation process. A "rolling" framework is sometimes used, and periodically updated and amended. In countries where external assistance forms an important part of the health system's resources, an important expansion of this approach to policy-making and implementation is represented by sector-wide approaches (SWAPs). The essence of SWAPs is that, under government leadership, a partnership of funding agencies agrees to work together in support of a clear set of policy directions, often sharing many of the implementation procedures, such as supervision, monitoring, reporting, accounting, and purchasing. Box 6.3 summarizes the development of SWAPs. Health planning thus shows signs of moving beyond investment programming and towards consensus statements on broad lines of policy and system development.

A policy framework should recognize all three health system goals and identify strategies to improve the attainment of each. Few countries have explicit policies on the overall goodness and fairness of the health system. Yet the need to combine these two values in governance can be traced far back in history (1). Box 6.4 describes the ancient *Hisba* system of stewardship in Islamic countries, highlighting both its ethical and economic purposes. Public statements about the desired balance among health outcomes, system responsiveness and fairness in financing are yet to be made anywhere. Policy should (and in partial

### Box 6.3 SWAPs: are they good for stewardship?

A sector-wide approach (SWAP) is a method of working that brings together governments, donors, and other stakeholders within any sector. It is characterized by a set of operating principles rather than a specific package of policies or activities. The approach involves movement over time under government leadership towards: broadening policy dialogue; developing a single sector policy (that addresses private and public sector issues) and a common, realistic expenditure programme; common monitoring arrangements; and more coordinated procedures for funding and procurement. Being engaged in a SWAP implies commitment to this direction of change, rather than

the comprehensive attainment of all these different elements from the start. It implies changes to the ways in which both governments and donor agencies operate, and in their required staff skills and systems.

This approach has begun to take root primarily in some of the most highly aid-dependent countries. It has been driven by both government and donor concerns about the results of historical approaches to development assistance, which have often involved a combination of 'social sector-blind' macroeconomic adjustment policies and 'sector-fragmenting' projects. Many of the countries are in Africa, for example, Burkina Faso, Ethiopia, Ghana, Mali, Mozambique, Senegal, Uganda, the United Republic of Tanzania, and

Zambia. The other cluster of countries discussing or actively engaging in a SWAP is in Asia: Bangladesh, Cambodia, and Viet Nam are examples.

The evolution of a SWAP takes time. In Ghana, before the Ministry of Health single sector programme was endorsed by donors, the country had already gone through 10 years of institutional development, 4 years of major policy/strategy work, 3 years of strengthening core management functions, 2 years of negotiations, planning and design, and 1 year of slippage and delays.<sup>1</sup>

Cambodia and Viet Nam are at the earliest stage of discussing sector policy with donors. In other countries, progress has been mostly towards developing and agreeing to

operate within a single sector policy and medium-term expenditure framework. Joint review missions have become a feature in some countries. Least progress has been made towards common financing and procurement arrangements.

SWAPs have the potential to support good stewardship. Walt and colleagues argue that SWAPs are perceived as capable of strengthening governments' ability to oversee the entire health system, develop policies and engage with stakeholders beyond the public sector.<sup>2</sup> But, most importantly, SWAPs depend on vision and leadership by national government.

<sup>1</sup> Smithson P. Cited in Foster M. *Lessons of experience from sector-wide approaches in health*. Geneva, World Health Organization, Strategies for Cooperation and Partnership, 1999 (unpublished paper).

<sup>2</sup> Walt G et al. Managing external resources in the health sector: are there lessons for SWAPs? *Health Policy and Planning*, 1999, 14(3): 273–284.

ways sometimes does) address the way in which the system's key functions are to be improved.

With respect to the provision of services, all providers should be recognized and their future contribution – greater in some cases, less in others – should be outlined. On financing, strategies to reduce dependence on out-of-pocket payments and to increase prepayment should be identified. Roles of the principal financing organizations – private and public, domestic and external – and of households should be recognized and their future directions determined. The machinery of stewardship, designed to regulate and monitor how these functions change in accordance with policy, should also be made explicit. This is likely to involve opportunities for consumer representatives to balance provider interests.

Danger exists when particular lines of policy, or whole reform strategies, become associated with a specific political party or minister of health. Regardless of whether the policy is good or bad, it becomes highly vulnerable. When that minister or party leaves office the policy dies, usually before it has either succeeded or failed, because the next minister or administration is seldom willing to work under the predecessor's banner. Rapid turnover of senior policy officials, and a politically charged environment, are both hazards to good stewardship (11). Establishing good stewardship can reduce exposure to “personality capture” of particular policy directions, by creating an informed constituency of stakeholder support, and ensuring that the interests, skills and knowledge needed to maintain a particular policy direction are widely distributed.

All remaining stewardship tasks concern the implementation of policy, as distinct from its formulation and promotion.

## SETTING THE RULES, ENSURING COMPLIANCE

Regulation is a widely recognized responsibility of health ministries and, in some countries, of social security agencies. It covers both the framing of the rules to govern the behaviour of actors in the health system, and ensuring compliance with them. In keeping with

### Box 6.4 Stewardship: the Hisba system in Islamic countries

The institution of Hisba was developed to carry out the function of stewardship in Islamic countries more than 1400 years ago. The *Hisba* system is a moral as well as a socioeconomic institution, whose *raison d'être* is to ordain good and forbid evil. The functions of the *muhtasib* (the head of *Hisba* system) can be classified into three categories: those relating to (the rights of) God; those relating to (the rights of) people; and those relating to both.

The second and third categories are related to community affairs and municipal administration. The

main foundation of *Hisba* was to promote new social norms and develop the required system to ensure the adherence of various sectors of society to these norms.

The first *muhtasib* in Islam was a woman called Al Shifa, appointed in Medina, the capital of the Islamic state, by the second calif, Omar ibn Al Khattab, almost 1450 years ago, and given authority to control the markets. Another woman called Samra bint Nuhayk was given a similar authority in Mecca, the second city, by the same calif.

The *muhtasib* could appoint technically qualified staff to investigate

the conduct of different crafts, trades and public services, including health services. The *muhtasib* received complaints from the public but could also order an investigation on his or her own initiative.

Medical services were also regulated by the Hisba system. Physicians and other health specialists had to pass professional examinations and possess the necessary equipment before being licensed. The *muhtasib* had to ensure compliance of practising physicians to moral and ethical norms, including equitable provision of services and protection of the public interest. In

the field of pharmaceutical services, technical publications were prepared, including monographs describing standards and specifications for various drugs as well as methods of quality assurance. The system also included inspections and enforcement mechanisms.

Like many other institutions, the Hisba system underwent drastic modification with the advent of western colonization: its functions were transformed into a number of secular departments and its moral content reduced.

Contributed by the World Health Organization Regional Office for the Eastern Mediterranean.

Source: Al-Shaykh al-Imam Ibn Taymiya. *Public duties in Islam: the institution of the Hisba*. Markfield, UK, The Islamic Foundation, 1985.

the policy-making and intelligence tasks, regulation has to encompass all health actions and actors, and not just those of the health ministry or the public sector. While the public health care system is often replete with regulations, few countries (with either high or low income) have developed adequate strategies to regulate the private financing and provision of health services. The rethinking of a consistent set of regulatory approaches to private providers and sources of finance, in line with national goals and priorities, is a top priority task in most countries.

*Regulation can either promote or restrict.* Since the private sector comprises many different players, national policy needs to distinguish carefully where to promote and where to restrict. A single position on the private sector is unlikely to be appropriate. In promotive terms, explicit incentives may be provided for private practice such as the sale of public assets, preferential loans, or donations of land. Tax incentives may be offered to promote private provision, with no or very little government regulation of providers' market behaviour. China re-legalized private practice in the 1980s and promoted joint public/private ventures in hospital ownership. Thailand's finance ministry offers tax incentives to private hospital investors.

At the other extreme, significant barriers to market entry have sometimes been created, such as a legal ban on private practice. This is still the case in Cuba and was previously in Ethiopia, Greece (for hospitals), Mozambique, the United Republic of Tanzania and several other countries. Between these extremes are policies that allow relatively free market entry, provide modest incentives, or have limited prerequisites for those wishing to enter the private market, including some standards for market behaviour and some level of oversight and enforcement.

Incentives for greater private sector opportunities in health are often sought by government agencies other than the health ministry, and by private investors themselves. Finance, trade, and development ministries often advocate greater private investment in health in line with overall economic liberalization strategies.

Promotive policies seem to work, contributing to growth in private finance and provision (12, 13). But they have also had serious side-effects: rising inequities, uneven quality of care, and inefficiency. The health ministry needs to know in advance what conditions it will require for such investments to contribute to the efficiency, quality, or equity goals of the health system, and how to defend the view that health is *not* just like all other sectors.

The harm caused by market abuses is difficult to remedy after the fact. The United States is probably the best-documented case of regulators trying to catch up with private health insurers (14). State governments have extensive laws, regulations and enforcement authority over private insurers in the USA to ensure fair competition, assure quality and generally protect consumers from fraudulent marketing. This regulatory framework took many years to develop and is still far from perfect: it does not guarantee insurance for everyone. Recent regulatory changes have improved access to, but not the affordability of, private insurance by small employers and individuals. Private employers have devised various ways of avoiding the rules, so as to come under the looser federal regulations. But the system prevents many of the worst abuses – financially unsound or unscrupulous insurers – and helps to ameliorate many market failures. Chile and South Africa have similar experiences in regulating private health insurance practice. South Africa has recently changed earlier regulations governing medical schemes to reduce risk selection and increase risk pooling (see Box 6.5).

Chile has been unable to establish explicit contractual obligations for private insurers or prohibit risk selection by these private companies, due to the political influence of insurers

and their clients. If there is a long delay between market entry and the enforcement of rules regarding market behaviour, experience suggests that the task of instituting those rules will become politically very difficult (15, 16).

A more moderate form of incentives for private sector involvement are represented by contracts between public purchasers and private providers. In Lebanon, for example, 90% of hospital beds are in the private sector and nongovernmental organizations provide ambulatory care to about 10–15% of the population, particularly to the poor. Out of necessity, the Ministry of Health contracts with almost all private hospitals for a predetermined number of beds to serve public patients (17). But the government does not use this regulatory tool to its advantage. Reimbursement policies allow unnecessary hospitalizations and overuse of services, which result in cost escalation; and private hospitals operate in a largely unregulated environment, which leads to uncontrolled investment. This in turn can lead to pressure for sustained public financial support, which will appear to justify further investment. Stewardship needs to ensure consistency in the incentive messages sent out by different levels of public policy.

**Regulation requires resources.** Regulatory oversight and contractual strategies entail high transaction costs for both government and providers or insurers, which may reduce the potential cost savings of these strategies. High levels of awareness of these costs accompanied the moves to separate the roles of purchasers and providers in the United Kingdom and New Zealand (18). Often, lack of commitment and funds hamper government capacity to carry out regulatory responsibilities, old as well as new. This suggests that capacity building in contracting skills and regulatory oversight is critically needed both via recruitment of skilled staff and through training and technical aid to existing staff.

### Box 6.5 South Africa: regulating the private insurance market to increase risk pooling

The government which came to power in 1994 after South Africa's first democratic elections found itself with a health sector which mirrored the inequalities existing in the wider society. A long-established and well-developed private health care industry accounted for 61% of health care financial resources, while providing for the needs of only the affluent 20% of the population. The vast majority of the population had to rely upon poorly distributed, underfunded and fragmented public services. Cost escalation in the private sector typically exceeded inflation during most of the late 1980s and 1990s. The private sector responded to this by limiting benefits, increasing co-payments and accelerating the exclusion of high-risk members from cover, thereby

heightening the problem of inequality.

The new government's response to these challenges was to enact new legislation for medical schemes to offer a minimum benefits package and increased risk pooling. The fundamental principles and objectives at the core of the Act are as follows.

- **Community rating.** For a given product or option, the only grounds on which premiums may be varied are family size and income. Risk or age rating are prohibited.
- **Guaranteed access.** No-one who can afford the community rated premium may be excluded on grounds of age or health status.
- **Increased risk pooling.** Caps on the permissible contributions and accumulations through individual medical savings accounts will en-

sure that a greater proportion of contributions flows into the common risk pool.

- **Promoting lifetime coverage.** Community rating and guaranteed access will be combined with premium penalties for those who choose only to take out cover later in life, to provide powerful incentives for affordable lifetime membership.
- **Prescribed minimum benefits.** Every medical scheme must guarantee to cover in full the cost of treating a specified list of conditions and procedures in public facilities, thus greatly decreasing the impact of "dumping" patients onto the state.

A committee of inquiry was appointed by the health minister during 1995. It set up a small technical

team to prepare new regulations for medical schemes. The team produced its first discussion document in 1996, and consulted widely with key stakeholders on its proposals. Discussion and debate continued until mid-1997, when a formal policy paper resulted.<sup>1</sup> After a period of intense, open debate during the legislative process, the new Medical Schemes Act and its accompanying Regulations came into force on 1 January 2000, three and a half years after the committee was formed. One important group will benefit immediately: HIV-positive members of medical schemes now have access to subsidized care, including drugs for opportunistic infections, whereas previously they were excluded or their entitlement was limited to very low benefit levels.

Contributed by T. Patrick Masobe, Department of Health, South Africa.

<sup>1</sup> *Reforming private health financing in South Africa: the quest for greater equity and efficiency.* Pretoria, Department of Health, 1997.

Shortcomings in staff skills or resources are often cited as the cause of outdated regulatory frameworks, or those which are not adequately enforced (4). Lack of legislative authority, too, is sometimes at fault. For example, in the late 1970s, Sri Lanka deregulated private practice by government doctors and liberalized the economy in general, which increased availability of capital (19). However, the health ministry failed to register effectively the growing number of private providers. It had no regulatory strategy, no staff responsible for private sector relations, and it lacked adequate legislative authority to take on many tasks. The only law on the books required registration of nursing homes, but not private clinics or doctors. A law has been pending since 1997 but has not yet been implemented. However, a new Ministry of Health unit for development and regulation of the private sector was set up in 1998.

In Egypt, most physicians work simultaneously for the government and in private practice. As a result, much of their work escapes oversight and regulation. Similar practice is widespread in Latin America. In India, mechanisms for monitoring, let alone regulating, the private sector have not kept pace with its expansion, despite concerns about quality of care. Health professionals are aware of practice-related laws but know that enforcement is weak or non-existent and that professional associations, which are nominally responsible for self-regulation, are also ineffective.

When public providers illegally use public facilities to provide special care to private patients, the public sector ends up subsidizing unofficial private practice. It is nearly impossible to completely prohibit private practice by health workers on the public payroll, but several steps can be taken to ensure that private practitioners compete on a fair basis and do not flourish by “moonlighting” at public expense (20, 21). Ensuring that patients, the public, and the media, as well as providers, know the rules is an important factor in regulating the public-private mix.

*Effective public services themselves can be a regulatory tool.* Developing effective public provision and financing systems becomes even more important if government policy seeks to restrict the development of a private health market, or when it lacks the resources to prevent undesirable market failures. The public sector must then respond to the changing needs of consumers, to the introduction of new medical technologies, and to reasonable expectations of health professionals. A strong public sector may even be a very good strategy for regulating private provision and for consumer protection, if it helps to keep the private sector more competitive in price and quality of service.

Too often, however, it is the public sector which is seen as uncompetitive in terms of quality and responsiveness, in spite of its free or subsidized services. If the public system deteriorates or does not continually improve, an unhealthy amount of resources and attention will be siphoned off trying to catch offenders in the “black market”, and growing under-the-table payments will undermine equity goals.

Rules rarely enforced are invitations for abuse. Stricter oversight and regulation of private sector providers and insurers is now on the policy agenda of many countries. But progress is slow if not impossible. This suggests that countries must not only consider the impact of the private sector on the public sector and develop the regulatory framework to limit deleterious effects, but must make a continuing commitment to enforce the rules by investing in the knowledge and skills of regulatory staff. A study in Sri Lanka concluded, “the slow response in the 1980s makes the regulatory task in the 1990s more difficult: uncoordinated and unmonitored private sector growth has created a market context which is bigger, more complex, and with more established provider and user interests” (19).

Professional self-regulation, as distinct from personal self-interest, supports good practice. In establishing a professional organization, health workers assume several of the basic tasks of stewardship – identifying and certifying members, sharing experience, and sometimes offering in-service training. Small amounts of financial support to such organizations can ensure that basic information needed on non-government providers, particularly in ambulatory care settings, is available to the ministry of health. In several East African countries where religious groups are important providers of health services, central, nongovernmental coordinating bodies already perform this role. National medical associations are common; associations of traditional practitioners also exist.

Recent reforms in the Netherlands demonstrate the difficult balancing act between stronger regulation to protect consumers and increase equity, and looser rules to allow more competition (see Box 6.6).

Developing countries have also implemented policies which help ensure that private actors work on behalf of the larger public good. In addition to the South African example in Box 6.5, Bangladesh's National Drug Policy, adopted in 1982, prohibits importation and sale of all non-essential drugs. As a result, about 1666 products that were judged ineffective or harmful were banned, while about 300 were approved for marketing. The government also oversees production quality of all manufacturers and provides training to drug retailers on rational drug use. "Through a combination of public sector oversight and private initiative, essential drugs have been placed within reach of large numbers of the population, [and there are] reasonable and stable drug prices for products ... produced locally" (22).

*Regulation requires dialogue.* In countries with stronger oversight of the private sector, governments for the most part place their regulatory structure at arms-length from the regulated private players. If they do not, the private sector can subvert the system through "regulatory capture", i.e. coopting regulators to make the regulations more favourable to them. But "arms-length" does not mean no communication. Dialogue between public policy-makers or regulators and private sector players is a critical factor in making such regulations work. Governments must not only see well for good stewardship, they must also listen. Groups that have both public and private representation provide valuable input into policy development and rule-writing by assessing how private sector players can contribute to public policy goals without compromising their ability to succeed in the market. The drawback of such processes is that they may slow the pace of reform. And even with strong oversight and regulation, private sector players can weaken the regulatory apparatus through political pressure.

In conclusion, the following important lessons for the development of regulatory frameworks for private health markets are clear.

### Box 6.6 Opening up the health insurance system in the Netherlands

The Netherlands' new health insurance system, authorized in 1990, for the first time required all private insurers to provide a comprehensive uniform benefits package. But it promoted competition by giving individuals a subsidy to help them buy compulsory health

insurance from competing insurers. Insurers receive risk-adjusted per capita payments by the government and a separate flat rate premium from each insured person. The more efficient the insurer, the lower the premium paid by the insured. Insurers were also allowed to negotiate

lower fees than officially approved provider fees, which was previously prohibited. As a result, private health insurers entered the market for the first time since 1941, and both insurers and providers became involved in quality improvement efforts, which became the focus for

competition among insurers rather than competition only on price.<sup>1</sup> But the new system made the goal of reducing health-related inequalities more difficult, as better-off individuals can prepay for more inclusive benefit packages.<sup>2</sup>

<sup>1</sup> Van de Ven W, Schut F. Should catastrophic risks be included in a regulated competitive insurance market? *Social Science and Medicine*, 1994, 39(1): 1459–1472.

<sup>2</sup> Saltman RB, Figueras J, Sakellariades C, eds. *Critical challenges for health care reform in Europe*. Buckingham, UK, Open University Press, 1998.

- Frameworks should be instituted *prior* to any significant planned expansion through economic incentives and forcefully implemented as soon as the private market starts to respond to incentives.
- Regulatory policies must be continually reviewed to ensure consistency with changing political scenes.
- Improving quality, increasing access to care, and promoting efficiency each require different regulatory tools.
- Regulators must strike a balance between avoiding regulatory capture by private interests and maintaining productive dialogue with them to ensure that regulatory frameworks are realistic.
- Where governments choose to restrict the activities of the private sector, they must ensure that the public sector responds effectively to the needs of consumers.

Governments must make a continuing commitment to enforce regulations and rules by investing in the knowledge and skills of regulatory staff to keep pace with market developments.

## EXERCISING INTELLIGENCE, SHARING KNOWLEDGE

Stewardship is about vision, intelligence and influence. Without a good understanding of what is happening in the entire health system, it is impossible for the ministry of health to develop strategies to influence the behaviour of the different interest groups in ways that support, or at least do not conflict with, the overall aims of health policy.

A good intelligence system in the sense of both information and understanding needs to be selective in the information it generates for decision-making at the top. But it must be drawn from grass-roots knowledge. Who are the principal service deliverers, and what challenges do they pose to health policy goals? Where are the main imbalances or bottlenecks in input production, and what policy options appear most suitable? Where are the major financing sources and what strategies will achieve greater and more equitable prepayment? What are the main uses of financing and what policies will ensure more efficient resource allocation?

Most health systems collect huge amounts of information that can clog the works. Such information may include accounts, personnel records, inventories, vehicle log books, activity reports (daily, by programme, department, ward, prescription and patient) at each health facility, and patient records. In many ministries of health, thousands of clerical hours each month are wasted in compiling information that is never used. As a general management rule, the amount of information passed up the system should be greatly reduced for each level.

For stewardship purposes, only periodic summaries, showing geographical or temporal variation, may be required. Information on the distribution and activity of public sector health inputs and on budgetary allocations may reveal important and unjustified variations. But of greater importance for stewardship are the missing pieces of information and analysis. Few low and middle income countries today have reliable information on the levels and sources of non-government finance or provision in the health system. As the national health accounts indicators in Annex Table 8 show, these are typically dominant in such countries. Little is known in most countries about peoples' expectations of the health system or about the structure of complex non-government provider markets. Without these data, assessments of responsiveness and fairness in financing, or of intermediate measures

such as service quality and accessibility, are impossible. Without the full picture, good stewardship cannot be practised.

*Intelligence requires resources.* Stewardship requires a different type of information and understanding from that required in the daily management of service delivery. Should the ministry of health collect it? There is no reason to assume that the resource and skill cost of stewardship intelligence is greater than that of traditional health management systems. Of course, new skills in the area of regulation, coordination and communication are needed. But the ministry of health may already have several advantages.

First, the dispersed national network of public sector health workers and managers provides skilled people for undertaking inventory or survey work. District level health workers can rapidly compile an initial register of non-government providers. Second, the ministry of health has the moral authority to license and accredit providers, so it can engage its staff in the assessment process. Third, health workers have frequent contact with the population and are well placed to ask people about public goals and personal expectations. So the ministry of health can be a formidable potential resource for better stewardship, beginning with its engagement for better intelligence on the entire system.

However, not all of the intelligence gathering, or sharing, will be best done by the ministry. Research institutes, university departments, nongovernmental organizations (23, 24) and local or international consulting firms may be able to undertake inventory and survey work more speedily and accurately. To manage them, the ministry will need to draw on skills in contract setting and oversight.

Stewardship also requires information for influencing behaviour and events. Information dissemination provides support, for instance, to both policy-making and regulation. It also allows the ministry to build a constituency of public support for health policy, and thus a defence against incompetent or corrupt practice by interest groups in the health system. It helps to achieve a public debate on policy directions that is based on reliable evidence. A strategy for disseminating technical information can also form part of a capacity-building

### Box 6.7 Responsiveness to patients' rights

Since the end of the 1970s there has been a slowly growing recognition of the rights of patients, such as respect for the dignity of the individual and for autonomy.

Rapid advances in medical and health sciences and in technology have hastened increases in patients' expectations: better-informed patients have begun to assert their rights in their dealings with professionals. To a growing extent, patients' rights are incorporated into statutory regulations: in laws on specific subjects, or in citizens' rights covering sectors broader than health care. Regulation may give patients direct legal rights in their relationships with health care providers, or may help to improve their position through

administrative health laws and hospital certification, for example. Self-regulation – voluntary arrangements in the form of professional codes or model contracts worked out in cooperation between consumers and health care providers' organizations – also have a role to play. Legislation opens new domains for self-regulation: framework laws on privacy and confidentiality, for example, may oblige institutions to elaborate their own guidelines for the protection of patients' data.

Three types of approaches can be distinguished in national legislation on patients' rights. Some countries have enacted a single comprehensive Law (e.g. San Marino in 1989, Finland and Uruguay in 1992, the Netherlands in 1994, Israel and

Lithuania in 1996, Argentina and Iceland in 1997, Denmark in 1998, and Norway in 1999). Other countries have integrated patients' rights into legislation regulating the health care system or into several health laws (e.g. Canada (New Brunswick) and Greece in 1992, France in 1992–94, Austria in 1993, Hong Kong in 1995, Belarus and Canada (Ontario) in 1996, Georgia and Guinea in 1997, and the USA in 1999). Charters on the rights of patients, which have varying status as national policy or are often embodied in the regulations of health care establishments, have been found more appropriate to the legal traditions of some countries, such as France (1974–95), Ireland (1991), the United Kingdom (1991–95) and Portugal (1997).

Informed consent, access to medical records, and the confidentiality of data are the classic rights of patients. New rules for the protection of personal data in medical data banks or automated hospital information systems are also being developed. In recent years the right to privacy has given rise to new individual concerns such as the right to be notified when personal data are first recorded in a data bank, the right to have inexact or incorrect data corrected or destroyed, and the right to be informed about the disclosure of information to third parties.

programme within the health system, and particularly within the ministry of health.

Information dissemination should focus on getting the most difficult tasks of stewardship into the open, both to inform and to consult. Priority setting in health, discussed in Chapter 3, has only recently been conducted as a public debate in a small number of countries. The debate is often noisy and confused because it lacks rules. The ministry's role is to clarify the rules: priority setting should take into account the burden of illness, the cost-effectiveness of available interventions, and the scale of existing action to address the problem. And it can listen to expressed preferences regarding the value basis of priority setting, as occurred in Sweden and Oregon, USA (25). The rights and obligations of different players can be clarified through dissemination strategy in ways which reinforce the concerns of policy. For example, in situations with prevalent informal charging for care, providers may at least be required to display publicly the full costs of procedures, and patients invited to register complaints where additional charging occurs.

Many countries have already taken steps to safeguard the rights of patients, as shown in Box 6.7. Even without legislation, the notion of patients' rights and providers' obligations can be promoted and given substance by active stewardship. Where particular practices and procedures are widely practised and known to be harmful, the ministry as a steward has a clear responsibility to combat these with public information. Pharmaceutical sales by unregistered sellers, the dangers of excessive antibiotic prescription and of non-compliance with recommended dosages should all be objects of public stewardship, with active support from information campaigns targeted at different actors – patients, the providers in question, and local health authorities. Box 6.8 illustrates how for one key input – pharmaceuticals – actions at different levels are needed.

### Box 6.8 Towards good stewardship – the case of pharmaceuticals

Most curative and many preventive health actions depend on medicines. However, medicines also involve powerful economic interests. In poor countries over 50% of household expenditure on health is spent on medicines: within government health budgets pharmaceuticals are usually the second largest item after wages. In industrialized countries drug costs are increasing by 8–12% per year, much faster than consumer prices. Many stakeholders are concerned with pharmaceuticals: manufacturers (both research-based and generic), consumer groups, professional associations, service providers of all types, donor agencies, and different departments of government.

The health system must make essential drugs available and affordable to all who need them, en-

sure that drugs are of good quality, and that they are used in a therapeutically sound and cost-effective way. The following are the core roles of central government to achieve these objectives:

- ensuring the quality of medicines through effective regulation including systems for market approval, quality assurance, licensing of professionals and inspection of facilities;
- ensuring the affordability and adequate financing of essential drugs for the poor and disadvantaged;
- procuring essential drugs for public sector providers, or establishing central tendering with prime vendor or delivery contracts for regional and lower levels;
- developing and supporting a national programme to promote rational and cost-effective drug

use by health workers and the public;

- coordinating the activities of all stakeholders through the development, implementation and monitoring of a national policy.

Good stewardship at the international level includes supporting governments in fulfilling these core roles. External support may also be useful in the following areas:

- nongovernmental organizations, professional and consumer networks, religious bodies, universities, and private providers need information support and management training;
- national pharmaceutical manufacturers need training, support and supervision in good manufacturing practice;
- regulations, training programmes and financial incentives are

needed to encourage rational drugs use in the private sector.

The international community must ensure that the overwhelming health problems of the world's poorest countries feature on the agenda of drug manufacturers; mechanisms such as the Global Alliance on Vaccine Initiatives and the Medicines for Malaria Venture are intended to do this.

In the technically and politically complex field of pharmaceuticals, external agencies may need guidance on the best types of support to give developing countries. For example, guidelines for good drug donation practice<sup>1</sup> are available to maximize the value of donated pharmaceuticals.

<sup>1</sup> *Guidelines for drug donations, 2nd ed.* Geneva, World Health Organization, 1999 (document WHO/EDM/PAR/99.4).

Broader information allowing comparisons of per capita health resources, and of health goal attainment by geographical area, are a way of spreading the stewards' concern about avoidable variations by creating public awareness. Without such awareness based on reliable information, government lacks an effective bulwark against incompetence and corruption in the form of personal or professional capture.

A recent study analysing initiatives in India, by the state governments of Delhi, Punjab and Rajasthan, to attract private investors into joint hospital ventures illustrates how the tasks of stewardship matter (26). All three schemes failed: no joint venture resulted. Different factors came into play in each situation, but the report identifies failure in each of the above tasks of stewardship in the overall explanation. It specifically identifies:

- inadequate policy on the role of the private sector by each state;
- insufficient consultation with relevant stakeholders, and absence of mechanisms for coordination among the parties concerned;
- absent, weak or inappropriate regulation machinery related to private providers;
- ineffective performance monitoring and information sharing arrangements, making public-private partnerships vulnerable to inefficiency and high cost.

Requisite skills for carrying out these tasks were found to be lacking in the health departments of all three states.

## STRATEGIES, ROLES AND RESOURCES: WHO SHOULD DO WHAT?

The previous sections discussed three basic tasks of stewardship and the principal role of the ministry of health in ensuring their implementation. This section considers *how* those tasks can be implemented, and what are the potential contributions of other groups and agencies to overall stewardship.

“Virtual” health systems, as described in Chapter 3, comprise many autonomous and semi-autonomous actors in different sectors of the economy, as well as those directly under the full authority of the ministry. The skills and strategies which have traditionally controlled public bureaucracies are inadequate for stewardship of contemporary health systems. Entrepreneurial, analytical and negotiating skills are needed to steward such systems. “Virtual” systems are held together by a shared policy vision and information, and by a variety of regulatory and incentive systems designed to reward goal achievement and punish capture, incompetence and fraud. An informed population of consumers helps in holding such a health system together.

Better stewardship requires an emphasis on *coordination, consultation and evidence-based communication* processes. For the ministry of health to understand the principal challenges to better performance it must have a full picture of what is happening. Initial engagement of other departments (education, finance, transport) may most effectively be done through government as a whole, rather than in bilateral approaches by the ministry of health, but the latter will need to provide evidence and continue the dialogue. Ministries of health can learn much from changing practice in other parts of government, where public roles have already greatly altered. And relevant international experience provides a major source of potential learning.

Ministries need to listen to a wider range of voices and to put the public case on health priorities and strategies forcefully and imaginatively. To ensure that the tasks of steward-

ship are carried out and delegated, the identity of all health actors should be known to the ministry of health, and regular lines of communication established. Special studies have sometimes been necessary (26) to assess the scale and content of private practice in health.

The ministry of health also needs communication capacity and strategies for ensuring that the media are aware of the health system's goals and progress or obstacles. Some ministries of health have offices responsible for private sector, media, and cross-sector liaison with other health players, and for consumer and public relations. In Thailand's experience, for example, skilful use of national media ensures that the Ministry of Public Health can amplify its own influence by judicious use of support (see Box 6.9).

Consultation is often a widely neglected part of the policy process, both in policy formulation and in implementation. A lack of consultation led to a public campaign of opposition by the British Medical Association to the reforms in Britain's national health service, introduced by the Thatcher administration in 1989 (27).

Kenya introduced its cost-sharing policy with substantial increases in user fees in December 1989. The press featured a number of hardship stories as a result of cost-sharing. The following August, a presidential announcement was made abandoning the policy. Fee policy was subsequently re-introduced in a phased way, beginning at specialist hospitals, with a much greater emphasis on staff training and familiarizing the public (28). Health system reforms in the United Republic of Tanzania and Zambia benefited from the Kenyan experience. They made great efforts to ensure that the reform programme was debated publicly, and that health workers were also involved in decisions about the reform process (29). Finland's system of democratically elected municipal health boards is cited as a good example of how to ensure citizens' participation and empowerment in health (30).

In many settings a sensible strategy to improve information for stewardship would be to begin with a review of key information needs for performance monitoring; develop strate-

### Box 6.9 Thailand: the role of the media in health system stewardship

Thailand is becoming a more open and responsive society. The 1997 Constitution foresees full democratic participation by the individual, community and civic society. The Public Organization Act (1999) grants government units autonomy, in close collaboration with civic society. Several public hospitals are being given autonomous status. Remaining public hospitals are setting up boards consisting of local lay members.

The Public Information Act (1998) further promotes transparency and social accountability through guaranteed citizen rights to government information. Amidst these reforms the media have played an important role in reflecting the public needs, and

have helped in shaping several key health policies. A Council of Journalists sets standards for ethical conduct and fosters balanced public information in the media. Regular public opinion polls help serve as an effective interface between the public and policy-makers.

The Ministry of Public Health has a long history of engaging support from many stakeholders, including the press and broadcasting media. Recent efforts have mobilized medical bodies and nongovernmental organizations to put sustained and public pressure on the government to promulgate two important laws, the Tobacco Products Control Act (1992) and the Non-smoker Health Protection Act (1992). This legal framework aims towards eventually achieving a smoke-free Thai society.

Traffic accidents are Thailand's leading cause of death. Intensive messages by radio and television during the highest traffic peaks have significantly reduced deaths and injuries in recent years. Other health activities such as physical fitness, healthy diet, and traditional medicines have been covered by radio channels providing evidence-based and balanced information. The media and nongovernmental organizations have set up HIV/AIDS counselling, and the Ministry of Public Health has set up a help line to provide counselling on stress and suicide prevention, as well as a telephone hot line aiming at consumer protection.

The media reflect public dissatisfaction with both public and private hospital care. At the same time the

Health Systems Research Institute (HSRI) coordinates a national forum on hospital quality improvement and accreditation and is pressing for an independent hospital accreditation body. HSRI also has a programme to guide journalists wanting to specialize in the health field.

Thus, Thailand's media play an important role in health system stewardship, as information providers and change agents, linking the general public, consumer groups, civic society, the research community, professional organizations and the government in improving health of the people in a participatory way.

gies for improving data collecting; review the core policy vision and messages; review existing organizational and incentive arrangements; and establish coordination and communication processes. A massive investment in management information systems will not, of itself, bring about better stewardship. Advocacy strategies, too, are needed to influence other branches of government and non-government health system players. The scope of regulation has to be broader, bringing in and giving voice to consumers, private providers, professional associations, and external assistance agencies.

An improved information base for policy creates a major strength for communication. On occasions this may require a higher profile by the ministry of health – in its dealings with the ministry of finance, or with donors, for instance. But the health ministry may get its messages across more forcefully when it uses other channels, such as the press, television and radio, academic institutions, and professional or consumer groups, to put its case. The ministry of health has to recognize all those primarily motivated by health gain – whether they are in the public or private sector – as its partners in the health system. Regular communication is one of the fibres which holds the system together.

The wide range of partners involved in a health system gives rise to an important question: who should do what?

Much of the preceding has been concerned with the role of the ministry of health. But the local context and particular issue determine who the stakeholders are – who stands to win or lose by a line of policy. Seeking the support of stakeholders is an important task for the ministry of health. The political feasibility of policy depends on: the power of the players; their position; the intensity of their commitment; and their numbers (31). As the agency responsible for formulating policy and steering its implementation, the ministry of health needs to bear this in mind.

Within the public sector, social security organizations and the education system are prominent among bodies whose activities affect health. The ministry of health can influence these either by dealing directly with them, or by working through higher political channels to ensure that health policies are supported, not contradicted by the practice of other parts of the public sector.

Where private sector activities are motivated by health gain, as for example in research and development in pharmaceuticals, medical technology, or motor vehicle safety, health ministries should at least ensure that their information and communication strategies include these partners. Where such inputs are internationally traded, regional and global organizations concerned with health should support the stewardship role of individual ministries of health by bringing governments, industry and consumer representatives together, promoting guidelines for good practice, and providing international information, monitoring and comparison.

Professional organizations can often play a much bigger role in self-regulation. With judicious support, ministries of health can assist professional bodies assume some of the burden of stewardship, such as licensing, credentials checking, and in-service training.

Consumer interests in health are weakly protected in countries at all levels of development. In countries such as Canada, New Zealand and Sweden, however, where public knowledge about health is taken seriously by government, numerically powerful and committed consumer groups have sprung up. Although they may oppose the ministry of health on some issues, on others the position of organized consumers will reinforce that of the ministry in dealings with input suppliers or professional groups. Modern communication strategies allow fast, easy access to health information in presentations suitable for non-

specialists: ministries should be energetic in making these resources available to the public.

External agencies, both public and nongovernmental, have special responsibilities in assisting stewardship. This report is directed to them and their expert advisers as well as to national policy-makers. External agencies have a dual mandate: they are accountable to their domestic chiefs and constituency as well as to governments in the developing countries in which they work. A focus on self-contained projects was a compromise which for many years made this dual mandate workable. Donors found projects an easy way to demonstrate their work to the home audience, and well-chosen projects also met a development priority need for the host country. The shift which began in the 1980s to more systemic support, through programmes and subsequently sector approaches, makes it much easier for external agencies to take a supportive role in government-led stewardship. Some donors now have a voice in the development of policy and strategy, and are abandoning their right to pick individual development projects in exchange for a fuller partnership with aid-receiving governments (32).

With their technical knowledge and resources, external agencies can ensure that the tasks of stewardship are recognized, and that the supporting investments in new skills needed to establish this function can be given priority. For stewardship is the irreducible core of public responsibility: government has to do this job and do it properly. Without stewardship, market failure and the exclusion of poorer consumers from access are ever-present dangers.

Donor agencies have a special responsibility not to make the stewardship role more difficult by acting in a semi-autonomous way. Donors – often numerous and anxious to ensure that their individual concerns are expressed in policy – can too easily find themselves at cross-purposes with each other and with government, compounding the difficulty of setting clear lines of policy (33). In this respect, the concept of sector-wide approaches offers a promising model. It puts government at the helm and establishes a dialogue on priorities, strategy and common implementation plans.

## WHAT ARE THE CHALLENGES?

Many countries are falling far short of their potential, and most are making inadequate efforts to achieve responsiveness and fairness in financing. There are serious shortcomings in the performance of one or more functions in virtually all countries.

These failings result in very large numbers of preventable deaths and disabilities in each country; in unnecessary suffering; in injustice, inequality and denial of basic rights of individuals. The impact is undoubtedly most severe on the poor, who are driven deeper into poverty by lack of financial protection against ill-health.

Within all systems there are countless highly skilled, dedicated people working at all levels to improve the health of their communities. There is little argument that health systems in general have already contributed enormously to better health for most of the global population during the 20th century. As the new century begins, they have the power and the potential to achieve further extraordinary improvements.

Unfortunately, health systems can also misuse their power and squander their potential. Poorly structured, badly led, inefficiently organized and inadequately funded health systems may do more harm than good.

The ultimate responsibility for the overall performance of a country's health system lies with government, which in turn should involve all sectors of society in its stewardship. The

Careful and responsible management of the well-being of the population – stewardship – is the very essence of good government. The health of the people is always a national priority: government responsibility for it is continuous and permanent.

Stricter oversight and regulation of private sector providers and insurers must be placed high on national policy agendas. Good policy needs to differentiate between providers (public or private) who are contributing to health goals, and those who are doing damage or having no effect, and encourage or sanction appropriately. Policies to change the balance between providers' autonomy and accountability need to be monitored closely in terms of their effect on health, responsiveness and the distribution of the financing burden.

Consumers need to be better informed about what is good and bad for their health, why not all of their expectations can be met, but that they still have rights which all providers should respect. But consumer interests in health are weakly protected in countries at all levels of development. The notion of "patients' rights" should be promoted and machinery established to investigate violations quickly and fairly.

The most obvious route to increased prepayment is by raising the level of public finance for health, but this is difficult if not impossible for poor nations. But governments could encourage different forms of prepayment – job-based, community-based, or provider-based – as part of a preparatory process of consolidating small pools into larger ones. Governments need to promote community rating, a common benefit package and portability of benefits among schemes, and to use public funds to pay for the inclusion of poor people in such schemes. Insurance schemes designed to expand membership among the poor are an attractive way to channel external assistance in health, alongside government revenue. Alert stewardship is needed to prevent the capture of such schemes by lower-risk, better-off groups.

Mechanisms are needed in most low and middle income countries to separate revenue collection from payment at the time of service utilization, thus allowing the great majority of finance for health to come through prepayment. More pooling of finance allows cross-subsidies from rich to poor and from healthy to sick. Risk pooling strategies in each country need to be designed to increase such cross-subsidies. Payments to service providers of all types need to be redesigned to encourage providers to focus on achieving health system goals through the provision of cost-effective interventions to people with common conditions amenable to prevention or care.

On an international level, the largely private pharmaceutical and vaccine research and development industry must be encouraged to address global health priorities, and not concentrate on "lifestyle" products for more affluent populations.

Serious simultaneous imbalances exist in many countries in terms of human and physical resources, technology and pharmaceuticals. Many countries have too few qualified health personnel, others have too many. Health system staff in many low income nations are inadequately trained, poorly paid, and work in crumbling, obsolete facilities with chronic shortages of equipment. One result is a "brain drain" of talented but demoralized professionals who either go abroad or move into private practice.

Overall, governments have too little of the necessary information to draw up effective strategies. National health accounts offer an unbiased and comprehensive framework from which overall situation analyses can be made, and trends monitored. They should be much more widely created and used.

## HOW TO IMPROVE PERFORMANCE

Stewardship is needed to achieve better health system performance. The following conclusions on stewardship apply in many industrialized countries as well as in low and middle income nations.

Stewardship of the health system is a government responsibility. To discharge it requires an inclusive, thought out policy vision which recognizes all principal players and assigns them roles. It uses a realistic resource scenario and focuses on achieving system goals. Intelligence requires a selective information system on key system functions and goal achievement, broken down into important population categories, such as income level, age, sex and ethnicity. Stewardship also calls for the ability to identify the principal policy challenges at any time, and to assess the options for dealing with them. Influence requires regulatory and advocacy strategies consistent with health system goals, and the capacity to implement them cost-effectively.

*Service provision.* Private provision of health services tends to be larger where countries' income levels are lower. But poorer countries seldom have clear lines of policy towards the private sector. They thus have important steps to take in recognizing the diverse forms of private provision and developing communications with the different groups of private providers.

In order to move towards higher quality care, a better information base on existing provision is commonly required. Local and national risk factors need to be understood. Information on numbers and types of providers is a basic – and often incompletely fulfilled – requirement. An understanding of provider market structure and utilization patterns is also needed, so that policy-makers know why this array of provision exists, as well as where it is growing. Information on the interventions offered and on major constraints on service implementation are also relevant to overall quality improvement.

An explicit, public process of priority setting should be undertaken to identify the contents of a benefit package which should be available to all, including those in private schemes, and which should reflect local disease priorities and cost-effectiveness, among other criteria. Rationing should take the form of excluding certain interventions from the benefit package, not leaving out any people. Supporting mechanisms – clinical protocols, registration, training, licensing and accreditation processes – need to be brought up to date and used. A regulatory strategy which distinguishes between the components of the private sector, and includes the promotion of self-regulation, needs to be developed. Aligning organizational structures and incentives with the overall objectives of policy is a task for stewardship, rather than one left only to service providers.

Monitoring is needed to assess behavioural change associated with decentralizing authority over resources and services, and the effects of different types of contractual relationships with public and private providers. Striking a balance between tight control and the independence needed to motivate providers is a delicate task, for which local – not textbook – solutions must be found. Experimentation and adaptation will be necessary in most settings. A supporting network for exchanging information will be necessary to create a “virtual health system” from a large set of semi-autonomous providers.

In middle income countries, where health service delivery is often segmented into parallel systems, quality-based competition among providers may be encouraged. A combination of public subsidy and regulated private providers, through extended insurance coverage

(Argentina, Colombia), and contracting directly to ministry providers (Brazil) has been implemented with some success. And in the high income economies, better regulation of private providers and greater attention to responsiveness (United Kingdom) and control of wastage due to over-prescribing, overuse of diagnostic technology and excessive interventions (France, Japan, the United States) are often needed.

**Resource generation.** Stewardship has to monitor several strategic balances and steer them in the right direction when they are out of equilibrium. A system of national health accounts (NHAs) provides the essential information base for monitoring the ratio of capital to recurrent expenditure, or of any one input to the total, and for observing trends. NHAs capture foreign as well as domestic, public as well as private inputs and usefully assemble data on physical quantities (numbers of nurses, CT scanners, district hospitals) as well as their costs. NHAs in some form now exist for most countries, but they are still often rudimentary and are not yet widely used as tools of stewardship.

NHA data allow the ministry of health to think critically about input purchases by all fundholders in the health system. The concept of strategic purchasing, introduced in Chapter 5, does not only apply to the purchase of health care services: it applies equally to the purchase of health system inputs. Where inputs such as trained personnel, diagnostic equipment, and vehicles are purchased directly with public funds the ministry of health has a direct responsibility to ensure that value for money is obtained – not only in terms of good prices, but also in ensuring that effective use is made of the items purchased.

Where health system inputs are purchased by other agencies (such as private insurers, providers, households or other public agencies) the ministry's stewardship role consists of using its regulatory and persuasive influence to ensure that these purchases improve, rather than worsen, the efficiency of the input mix. This does not, however, entail comprehensive central planning and programming. The role of stewardship in systems with a great deal of decentralized spending authority is to set the rules, rather than to adjudicate every decision. In Brazil, rules for allocating funds to states, prices for services, and reviews of major investment decisions have been put into practice (34). The central ministry may have to decide on major capital decisions, such as tertiary hospitals or medical schools. But regional and district health authorities should be entrusted with the larger number of lower-level purchasing decisions, using guidelines, criteria and procedures promoted by central government.

Ensuring a healthy balance between capital and recurrent spending in the health system requires analysis of both public and private spending trends and a consideration of both domestic and foreign funds. The budgetary information usually available to the ministry of health tells only part of the story. A clear policy framework, incentives, regulation and public information need to be brought to bear on important capital decisions in the entire system to counter ad hoc decisions and political influence.

In the field of human resources, similar combinations of strategy have had some success in tackling the geographical imbalances common within countries. In general, the content of training needs to be reassessed in relation to workers' actual job content, and overall supply often needs to be adjusted to meet employment opportunities. In countries such as China where the social return to medical training is negative, training institutions are being considered for privatization or closure. Certainly, public subsidies for training institutions often need to be reconsidered in the light of strategic purchasing. Re-balancing the intake levels of different training facilities is often possible without closure, and might free resources which could be used to retrain clearly surplus health workers (for example, specialist doctors in Egypt) in scarcer skills.

Stewardship of pharmaceuticals and vaccine inputs consists, at international level, of influencing the largely private research and development industry to address global health priorities. At national level the key tasks are to ensure cost-effective purchasing and quality control, rational prescribing, and that consumers are well informed. Health financing strategy also needs to ensure that poor people, in particular, get the drugs they need without financial barriers at the time they are sick.

Major equipment purchases are an easy way for the health system to waste resources, when they are underused, yield little health gain, and use up staff time and recurrent budget. They are also difficult to control. All countries need access to technology assessment information, though they do not necessarily need to produce this themselves. The stewardship role lies in ensuring that criteria for technology purchase in the public sector (which all countries need) are adhered to, and that the private sector does not receive incentives or public subsidy, including the subsidy inherent in being able to sell the services of that equipment to government, for its technology purchases unless these further the aim of national policy.

Providers frequently mobilize public support or subscriptions for technology purchase, and stewardship has to ensure that consumers understand why technology purchases have to be rationed like other services. The public case may be helped by identifying the opportunity cost of additional technology in terms of other needed services.

**Health system financing.** In all settings, very high levels of fairly distributed prepayment, and strategic purchasing of health interventions are desirable. Implementation strategies, however, are much more specific to each country's situation. Poor countries face the greatest challenge: most payment for health care is made at the time people are sick and using the health system. This is particularly true for the poorest people, who are unlikely to have any prepaid health insurance and who are frequently unable to benefit from subsidized services. Out-of-pocket payment for care, particularly by the poor, should not be relied on as a long-term source of health system finance.

Perhaps the most obvious route to increased prepayment is by raising the level of public finance for health, but two immediate obstacles appear. The poorest countries as a group manage to raise less, in public revenue, as a percentage of national income than middle and upper income countries. Indeed, this lack of institutional capacity is a facet of their poverty. And ministries of finance in poor countries, often aware that the existing health system is performing poorly, are sceptical of its claims on public revenues. Where there is no feasible organizational arrangement to boost prepayment levels, both donors and governments should explore ways of building enabling mechanisms for the development or consolidation of large risk pools. Insurance schemes designed to expand membership among the poor offer a path for government – with external funding partners – to a rapid improvement in the health of the most vulnerable.

In middle income countries substantial mandatory, income and risk-based schemes often coexist. The policy route to a fair prepaid system lies through strengthening such schemes, again ensuring increased public funding for the inclusion of poor people. Expansion of the beneficiary base through subsidies and merger of pre-existing schemes was how national coverage grew from small-scale schemes in Germany, Japan and the Republic of Korea.

Although most industrialized countries already have very high levels of prepayment, some of these strategies are also relevant to them. For its income level, the United States has an unusually high proportion of its population without health insurance protection: a

combination of the above strategies will be necessary if the level and fairness of financial protection is to be substantially improved in the decade ahead.

To ensure that prepaid finance obtains the best possible value for money, strategic purchasing needs to replace much of the traditional machinery linking budget holders to service providers. Budget holders will no longer be passive financial intermediaries. Strategic purchasing means ensuring a coherent set of incentives for providers, whether public or private, to encourage them to offer priority interventions efficiently. Selective contracting and the use of several payment mechanisms are needed to set incentives for better responsiveness and improved health outcomes.

This report has broken new ground in presenting for the first time an overall index of national health systems' attainment, and an index of performance relative to potential. These are based on the fundamental goals of *good health*, *responsiveness to people's expectations* (where both level and distribution matter for each of these goals), and *fairness of contribution to financing the health system*. Achieving these goals depends on the effectiveness of four main functions: *service provision*, *resource generation*, *financing*, and *stewardship*.

The preliminary ranking of countries in terms of their health system performance is revealing. It suggests that, at very low levels of health expenditure, performance is both systematically worse and much more varied than at high spending levels, even when performance is judged relative to a country's human resources and how much is spent on health. Clearly the countries with limited resources and severe health problems present the greatest needs: to understand why health systems do not achieve as much as it seems they might, and to help them attain their potential. The findings reported here also show that while much achievement – particularly for the level of health and some aspects of responsiveness – depends greatly on how much a system spends, it is possible to achieve considerable health equality, respect for persons, and financial fairness even at low resource levels. Some systems achieve much more than others in these important respects.

Much more work lies ahead for all concerned to improve the concepts and generate the data on national health system performance. A widespread refocusing of policy is strongly suggested.

Service delivery, resource mix, health financing and, above all, stewardship all matter greatly. The better performance of these four common functions makes substantial gains in goal achievement possible in countries at all levels of development. The poor will be the principal beneficiaries.

## REFERENCES

1. Saltman RB, Ferroussier-Davis O. On the concept of stewardship in health policy. *Bulletin of the World Health Organization*, 2000, 78(6) (in press).
2. Osborn D, Gaebler T. *Reinventing government*. Reading, MA, Addison Wesley, 1993.
3. Saltman RB, Figueras J, Sakellarides C, eds. *Critical challenges for health care reform in Europe*. Buckingham, UK, Open University Press, 1998 (State of Health Series).
4. Bossert T et al. Transformation of ministries of health in the era of health reform: the case of Colombia. *Health Policy and Planning*, 1998, 13(1): 59–77.
5. Ogunbekun I et al. Private health care in Nigeria: walking the tightrope. *Health Policy and Planning*, 1999, 14(2): 174–181.
6. Bennett S et al. *Health insurance schemes for people outside formal sector employment*. Geneva, World Health Organization, 1998 (unpublished document WHO/ARA/CC/98.1).

7. Killingsworth J et al. Unofficial fees in Bangladesh: price, equity and institutional issues. *Health Policy and Planning*, 1999, 14(2): 152–163.
8. Halik J. Respecting patients' rights in hospitals in Poland. In: *Health sector reform in Central and Eastern Europe: current trends and priority research: a FICOSSER research meeting, Velingrad, Bulgaria, 2–3 October 1998*. Warsaw, National Centre for Health System Management, 1999.
9. Schuster M et al. How good is the quality of good health care in the United States? *The Milbank Quarterly*, 1998, 76(4): 517–563.
10. Bennett S, McPake B, Mills A, eds. *Private health providers in developing countries: serving the public interest?* London, Zed Books, 1996.
11. Salinas H, Lenz R. *Las no reformas de salud en Latinoamérica: razones que explican su fracaso* [Non-reforms of health in Latin America: reasons explaining their failure]. Santiago de Chile, 1999 (in Spanish).
12. Nittayaramphong S, Tangcharoensathien V. Thailand: private health care out of control? *Health Policy and Planning*, 1994, 9(1):31–40.
13. Turshen M. *Privatizing health services in Africa*. New Brunswick, New Jersey, Rutgers University Press, 1999.
14. Chollet DJ, Lewis M. Private insurance: principles and practice. In: Schieber G, ed. *Innovations in health care financing*. Washington, DC, The World Bank, 1997 (World Bank Discussion Paper No. 365).
15. Musgrove P. *Public and private roles in health: theory and financing patterns*. Washington, DC, The World Bank, 1996 (World Bank Discussion Paper No. 339).
16. Hsiao WC. Abnormal economics in the health sector. *Health Policy*, 1995, 32:125–139.
17. Van Lerberghe W et al. Reform follows failure: unregulated private care in Lebanon. *Health Policy and Planning*, 1997, 12(4): 296–311.
18. Borren P, Maynard A. The market reform of the New Zealand health care system: searching for the Holy Grail in the Antipodes. *Health Policy*, 1994, 27(3): 233–252.
19. Russel S, Attanayake N. *Sri Lanka: reforming the health sector. Does government have the capacity?* Birmingham, UK, University of Birmingham, 1997.
20. Bennet S, Ngalande-Banda E. *Public and private roles in health. A review and analysis of experience in sub-Saharan Africa*. Geneva, World Health Organization, 1994 (Current Concerns, ARA Paper No. 6, document WHO/ARA/CC/97.6).
21. Broomberg J. *Health care markets for export? Lessons for developing countries from European and American experience*. London School of Hygiene and Tropical Medicine, Department of Public Health and Policy, 1994.
22. WHO Action Programme on Essential Drugs. *Public-private roles in the pharmaceutical sector*. Geneva, World Health Organization, 1995 (unpublished discussion paper).
23. Zurita B et al. *Structural pluralism as a tool for equity, quality and efficiency in healthcare in Mexico: the role of FUNSALUD*. 2000 (unpublished paper).
24. Smithson P. Quarts into pint jugs? The financial viability of health sector investment in low income countries. *Health Policy and Planning*, 1995, 10(Suppl.): 6–16.
25. Ham C. Priority setting in health care: learning from international experience. *Health Policy*, 1997, 42(1): 49–66.
26. Bhat R. *Public-private partnerships in the health sector: the Indian experience*. 2000 (unpublished paper).
27. Robinson R, Le Grand J, eds. *Evaluating the NHS reforms*. London, King's Fund Institute, 1994.
28. Quick JD, Musau SN. *Impact of cost sharing in Kenya: 1989–1993. Effects of the Ministry of Health Facility Improvement Fund on revenue generation, recurrent expenditures, quality of care, and utilization patterns*. Nairobi, Management Sciences for Health, 1994.
29. Kalumba K. *Towards an equity-oriented policy of decentralization in health systems under conditions of turbulence: the case of Zambia*. Geneva, World Health Organization, 1997 (document WHO/ARA/97.2).
30. Calnan M, Halik J, Sabbat J. Citizen participation and patient choice in health reform. In: Saltman RB, Figueras J, Sakellarides C, eds. *Critical challenges for health care reform in Europe*. Buckingham, UK, Open University Press, 1998 (State of Health Series).
31. Reich M. In: *Diagnostic approaches to assessing strategies, weaknesses and change of health systems*. Washington, DC, Economic Development Institute, The World Bank, 1998 (Flagship Module 2).
32. Hay R. *International aid: economics and charity*. Oxford, Oxford Policy Institute, 2000 (Oxford Policy Brief No. 1).
33. Walt G et al. Health sector development: from aid coordination to resource management. *Health Policy and Planning*, 1999, 14(3): 207–218.
34. *Brazil: social spending in selected states*. Washington, DC, The World Bank, 1999 (World Bank Report BR-17763: Chapter 3).