

WEEK 8 HEALTH SYSTEM REFORM AND THE DEMAND FOR HEALTH CARE

Despite enormous variation in per capita spending, health systems world-wide have undergone a series reforms in the last thirty years in response to claims about the rising costs of health care and the need to contain them. Market reforms have been widely advocated as the preferred method of cost containment on the ground that competitive markets improve cost efficiency. Reforms that target the 'demand' for health are known as 'demand-side' reforms. These reforms and the cost containment thesis are examined in this lecture.

Objectives/learning outcomes

Students will be able to:

Recognise and understand concepts of demand-side reform.

Critically examine demand-side policies.

Seminar: Demand-side health system reforms: what are they and how would you critically assess them?

Class discussion. Select a health system from the health systems in transition database (<http://www.euro.who.int/en/who-we-are/partners/observatory/health-systems-in-transition-hit-series>) and using the summaries identify mechanisms that have been introduced to combat demand-side cost inflation.

Describe the objectives and mechanism. How would you critically examine the public health effects of the policy?

Set reading

Cutler D (2002) Equality, efficiency, and market fundamentals: the dynamics of international medical care reform. *Journal of Economic Literature*; 40: 881-906.

Josep Figueras, Martin McKee, Suszy Lessof, Antonio Duran, Nata Menabde (eds) (2008) *Health systems, health and wealth: Assessing the case for investing in health systems*. Copenhagen: WHO. CHAPTER 4 ONLY.

Lecture summary

Introduction

Despite enormous variation in per capita spending, health systems world-wide have undergone a series reforms in the last thirty years in response to claims about the rising costs of health care and the need to contain them. For example, as early as 1996 a WHO review of European reform strategies noted:

'Health policy in Europe over the last two decades has been increasingly bedevilled by the growing cost of care. The aging of the population associated with higher levels of chronic disease and disability, the increased availability of new treatments and technologies, and rising

public expectations have exerted an upward pressure on overall health related expenditures.’ (WHO 1996, p.2)

Market reforms have been widely advocated as the preferred method of cost containment on the ground that markets improve cost efficiency. Cutler describes three main reform periods culminating in the widespread application of market or managed care mechanisms – the use of financial incentives to influence the supply and demand for health care (Cutler, 2002)

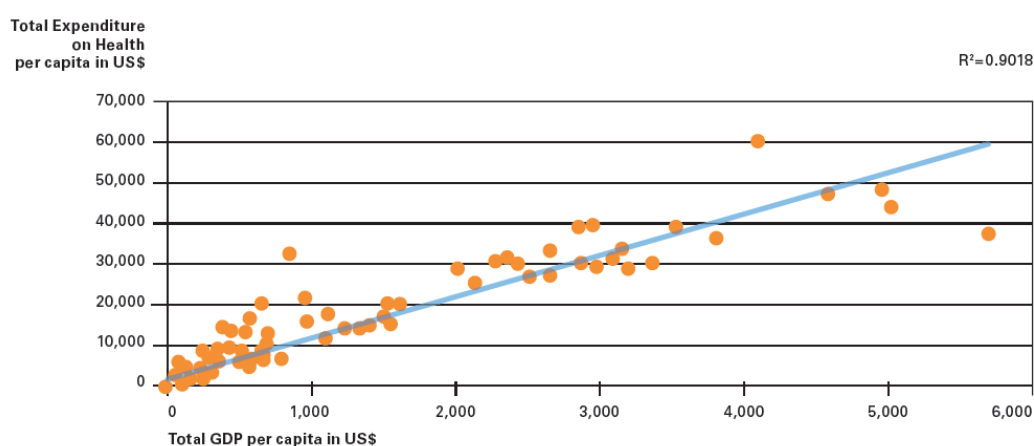
Privatization of management, funding and provision have been the main components of a reform movement that has relied heavily on the proposition that health care costs in developed and less developed countries are unaffordable. In less developed countries the World Bank and the WHO advise on health system reform in resource-poor contexts, in many cases advocating reforms similar to those in developed countries.

The reform agenda is intensely political because it involves reductions in state intervention and challenges redistributive policy. For example, Roberts et al (2004, pp.16-17) assert: ‘The global turn to the market has brought a trend toward diminished social solidarity and a parallel turn against government action in many countries.’ Though contestable, the statement underlines the reform movement’s major implications for redistribution. The conflict between market and planned redistribution is a recurring theme in the literature.

What are the determinants of total health care spending?

According to Donaldson (2005), the literature shows that income is the main determinant of the proportion of GDP a country spends on health care. The richer a country the greater proportion of GDP devoted to health spending.

WEALTH AND HEALTH EXPENDITURE ARE CORRELATED 2003



Source: WHO

Levels of expenditure may also vary according to mix of financing. In richer countries with high levels of public finance, there may be tighter control over

budgets. Countries lacking centralized control may have higher expenditure levels.

General trends in health system costs are discussed in Musgrove et al (2002) Huber (1999), Reinhardt et al (2002). Musgrove analysed data for 191 WHO member countries. He encountered serious data problems. He began with an analysis of total health spending relative to GDP and found:

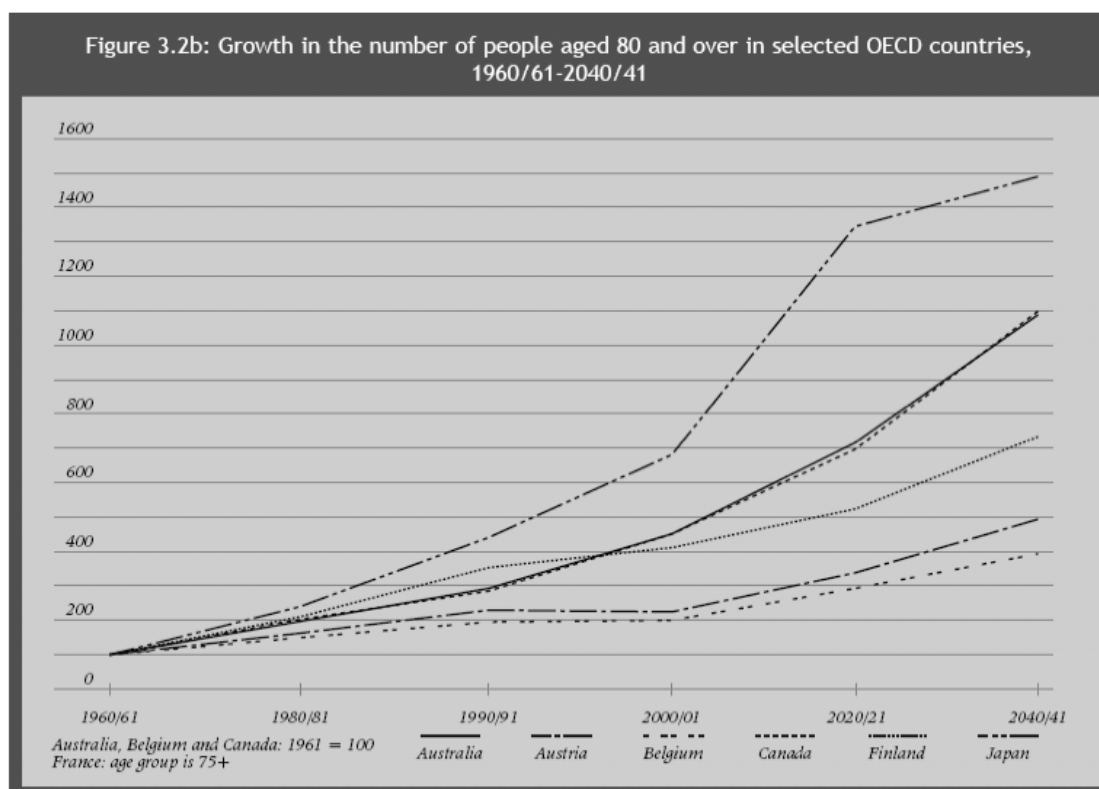
- Total health spending rises with GDP from an average of 2–3% of GDP at low incomes (<US\$ 1000 per capita) to 8–9% at high incomes (>US\$ 7000)
- *But that there is substantial variation around these averages in both low and high income countries. (Emphasis added)*

It is also claimed that, even where there are central controls, health care costs are rising at a faster rate than national incomes. This ‘cost inflation’ is usually attributed to three main factors:

- An ageing population (the ‘demographic time bomb’)
- Technological change
- Consumer demand

Although these claims are disputed they have heavily influenced health system proposals, which are often directed towards cost containment.

Example of disputed cost projections – ageing



Source: Replies to the OECD questionnaire on the care of the frail elderly, 1992

The OECD figure (above) estimated the growth in the number of people aged 80 and over in selected countries measured from a 1961 baseline. There is growth to 2040 in all countries surveyed. In the case of Japan a huge increase of 1300 per cent in 80 years is forecast. These growth rates are attributable to the common phenomenon throughout the developed world of a transition from high to low rates of mortality and fertility.

'Northern European countries have already a relatively aged population and the rate of growth is slower. Over the period 1960-2040 (or one average life span) the population aged 80 or over is expected to multiply by about three times in most northern European countries, rising by over 400% in Switzerland and over 600% in Finland. Even this rate of growth is dwarfed by that anticipated in the non-European industrialised countries, which is projected at a minimum of around 500% in New Zealand, over 800% in the United States, over 900% in Australia and Canada and over 1,300% in Japan. The relevant rate of increase in the United Kingdom is estimated at around 250%, and we are roughly half way along that curve now. Figures 3.2a, 3.2b and 3.2c show the OECD countries. The UK is near the bottom and Japan, Australia and Canada are at the top.' (Research vol. 1, chapter 1: 24-26)

The proposition is that health costs rise because a growing proportion of the population is elderly. But does increased life expectancy involve greater use of health services?

The suggestion was firmly rejected in a study for a royal commission on long term care carried out for the British government. (See Mahal and Berman (2001) and Research Volume 1, Part 1, Chapter 1 of Royal Commission on Long Term Care, & Executive summary, main report). The authors concluded that, even though 'the population aged 80 or over is growing rapidly and appears likely to continue to do so', the UK was not on the verge of a "demographic timebomb" as far as long-term care is concerned and as a result of this, the costs of care will be affordable.'

The authors reasoned that for health service planning 'it is not just numbers of people that matter. The health of older people has also to be considered.' The cost impact had therefore to be considered in the light of the need for health care, that is, rates of morbidity. The authors therefore argued that health expectancy was a more sensitive indicator of cost than life expectancy. Health expectancy is a measure of 'the numbers of years of life that will be free of chronic illness (which limits a person's ability to carry out tasks like cleaning or shopping) or severe disability (which prevents a person from looking after their personal needs such as feeding, going to the toilet, bathing or which inhibits their mobility)... Whereas life expectancy is the total number of years a person may

expect to remain alive, health expectancy adjusts this to reflect departures from good health during that expected lifetime.’

The Commission’s model projected public and private spending up to 2051. The base case showed a marginal increase in total spending on long term care costs expressed as a proportion of GDP (from 1.6% to 1.9%), that is, a cost increase from demographic change of 0.3% of GDP.

Table 3.5: Base case projections of long-term care costs for elderly people in the UK (1995/96 prices)

	1995 £ billions	2010 £ billions	2021 £ billions	2031 £ billions	2051 £ billions
NHS Continuing Care (1)	2.6	3.5	4.9	7.0	10.9
PSS net (2)	4.5	5.5	7.2	10.1	16.1
Private expenditure (3)	4.0	5.7	7.8	10.9	18.3
Total	11.1	14.7	19.9	28.0	45.3
% increase from 1995		32%	79%	152%	308%
GDP (total as percentage of GDP)	700 (1.6%)	980 (1.5%)	1,250 (1.6%)	1,560 (1.8%)	2,440 (1.9%)

Source: PSSRU and Royal 1995/96 PricesCommission

These forecasts are apparently not supported by data from the EU. According to Figueras et al (2008): ‘A European Commission (EC) study (26) based on the EU Member States belonging to the EU before January 2007 (EU25) predicts that the combined effects of health and long-term care will account for an additional 1.1–3.8% of European GDP...’ Nevertheless, the authors add that the figures should be treated with some caution because not enough is known and utilization rates are hard to predict.

Cost containment policy

Cost containment policy can focus on reducing the state’s responsibilities for health care (demand-side reform) and/or on increasing the efficiency of services (supply-side reform). Demand side reforms often involve changing the source of health system funding. Supply side reforms involve institutional and ownership changes believed to improve cost efficiency. Mossialos and Le Grand (1999, pp.62ff) summarise the main types of cost containment measure:

- budget shifting (co-payments, treatment restrictions, public budget shifting)
- budget setting (administrative controls)
- direct and indirect controls (financial incentives or managed care)

Demand side reform

Demand-side policies aim to reduce costs by eliminating 'unnecessary expenditure'. However, the question of what is or is not 'necessary' is neither clear cut nor easily determined. Unnecessary expenditure can be defined in three quite different ways. First, it can be defined as expenditure that is additional to an 'essential health care package'. This approach involves service targeting and rests on ideas about the practicability of providing public health care on a selective basis, that is, for a limited range of services. Secondly, unnecessary expenditure can be defined as that which is publicly financed but received by those who can afford to pay for it¹. This approach leads to user charge policies and depends upon the practicability of providing waivers for the poor. Third, it can be defined as expenditure on care that is not medically necessary.

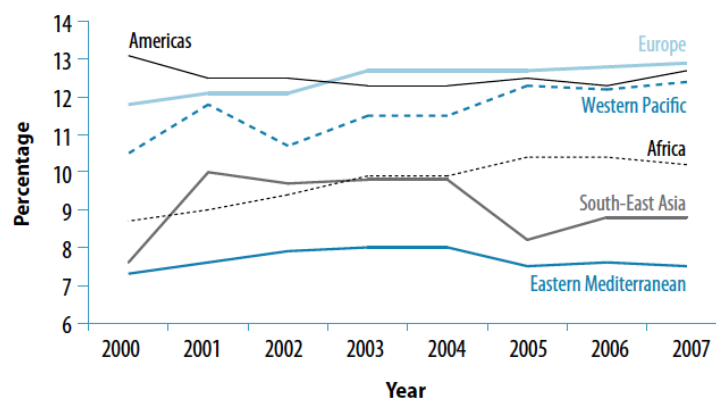
Demand side policy has been extensively influenced by research suggesting that insurance leads to over-consumption.

Researchers from the 1970s Rand Health Insurance Experiment claimed to show that insurance leads to 'moral hazard' (over-consumption) because prepayment and pooling insulate consumers from the full economic cost of their consumption decisions. The policy rationales of demand-side control draw on economic beliefs about the effects of insurance on demand and on ideas about the role of private insurance and managed care in addressing these effects, that is, in avoiding moral hazard. Managed care, which can operate through public or private insurance, aims to address moral hazard by shifting responsibility for the financial risks of care on to patients (among others).

Demand side policies are motivated by different considerations in the developed and developing world. In an earlier lecture I referred to arguments about the pressures of universal care on developed country public budgets and the extent to which health spending accounts for a growing proportion of the public budget. In developing countries attention often focuses on the inadequacy of the government's total budget.

¹ For the argument that universal benefits flow largely to the better off see Gwatkin D (2003) Free Government Health Services: Are they the best way to reach the poor?

Fig. 2.1. Government expenditure on health as a percentage of total government expenditures by WHO region, 2000–2007^a



^a These are unweighted averages. Government health expenditure includes health spending by all government ministries and all levels of government. It also includes spending from compulsory social health insurance contributions.

Source: (4).

See summary of issues in Schieber et al, Financing health systems in the 21st century, and in particular the problems of raising revenue in poorer countries:

‘In 2001, high-income countries spent an average of 7.7 percent of their GDP on health (country weighted), MICs spent 5.8 percent, and LICs spent 4.7 percent. Even though a clear upward trend between a country’s income level and the level of public and total health spending is apparent in terms of both absolute spending and share of GDP, spending for any given income level varies a great deal, particularly at lower income levels. The composition of health spending also exhibits major differences. As incomes increase, both private and out-of-pocket shares of total health spending decrease. In LICs, private and out-of-pocket spending and external assistance account for the bulk of all health spending. As countries move up the income scale, public spending predominates and both out-of-pocket spending and external assistance decrease drastically.’

There is also wide variation in revenue raising patterns in developing and developed countries. Developed country governments collect about 40% of GDP in revenues; in poorer countries the range is from 4% (Myanmar) to 36% (Lesotho). Crucially, the IMF, which as we have seen has substantial policy authority in poorer countries, has in the past advocated a norm for government spending of no more than 30% of GDP on the theoretical ground that public spending ‘crowds out’ private spending and as a result impedes economic growth. This norm has provided a separate rationale for international policy on user charges. That is, private finance has been advocated as a means of restructuring government spending as well as a means of containing unnecessary demand.

User fees were linked to these fiscal constraints in debates from the 1980s. For example, according to UNRISD: ‘it was argued that global competition called for changes in tax policies and the need to reduce “social wages” represented by social transfers. Not surprisingly, many of the debates on targeting in the 1980s revolved around restricting public spending so as to allow tax cuts, especially on

traded goods, and remove other taxes presumed to be “distortionary” and, therefore, the cause of poor export performance.’

The basic fiscal argument for targeting in developing countries is as follows: ‘in face of limited fiscal resources, it is better to target the resources to the “deserving poor”. Governments are presented as if they were confronted with an exogenously given fiscal constraint and are enjoined to do their best under the circumstances.’ (UNRISD)

Demand-side policies

1. Managed care

Managed care involves devolving financial risks through the use of financial incentives and reimbursement systems to change the behaviour of providers, doctors and patients (Donaldson and Gerard, 2004: 62; Morris et al., 2005; Liu, 2003). Erdmann (2001) gives an account of the adoption of managed care by European health systems because of its potential to limit the health care package: *‘From a European point of view, a striking characteristic of these new MCOs is their active and constraining influence on the extent and type of health care offered and provided to their subscribers.’* See also Kutzin (2003, 14) for a summary.

Managed care reforms are frequently associated with increase in the direct costs borne by patients. For example, World Bank policy has been to increase the role played by user charges in funding health systems (World Development Report, 1993).

Control is exercised over demand and supply by shifting responsibility for the financial of risks of care. Risk is devolved through contracts between insurers and providers. I quote Erdmann’s account (2001) of the adoption of managed care by European health systems:

In the 1970s the USA began adopting the managed care or HMO model to improve cost efficiency in health services. These plans ‘served as gatekeepers to both hospital and specialist services. [...] At the same time, insurance companies and other for-profit health care entities entered into the direct provision of services through this type of organization. [...] By 1998 this type of plan insured more than 30% of the population...’

There were fundamental changes in provision following vertical integration of those providing care and those financing it: *‘From a European point of view, a striking characteristic of these new MCOs is their active and constraining influence on the extent and type of health care offered and provided to their subscribers.’* These arrangements are associated with a system of user charges, co-payments or deductibles, all of which are demand side controls.

2. Targeting and rationing

Rationing can be used as an explicit device for limiting the availability of less cost effective treatments. The State of Oregon in the USA attempted in the 1990s to introduce rationing of this type. The attempt had to be abandoned because of the

public controversy it caused. (Maxwell RJ (2005) Why rationing is on the agenda. *British Medical Bulletin*; 51: 761-768. Provides a general argument for the need for rationing).

Known alternatively as service targeting, this approach is associated with efforts to concentrate public provision on services that are either a) measurably more cost efficient (for example, NICE-type exercises and the Global Burden of Disease, both of which involve economic evaluation of interventions, or b) essential to the poor (for example, selective primary care (as distinct from the primary health care model propounded at Alma Ata)). These 'essential' health care packages or vertical programmes are contrasted with universal, comprehensive care.

A second approach is to provide budgets within which clinical staff prioritize. Resource allocation formula can be used to apportion money to under-resourced areas. As we have seen, in poorer countries there can be difficulties in obtaining sufficient data for rational resource planning. Anne Mills et al (*Strengthening health systems in Schieber et al*) provides examples of proxy measures for under-resourced areas. Resource allocation techniques of this type were discussed in the introductory course.

3. Cost-sharing

Cost sharing refers to private or out-of-pocket (OOP) payments for health care services. There are several types of OOP.

In the private insurance and managed care market OOP takes the form of deductibles or copayment. Deductibles refer to the portion of an insurance claim that is not covered by the insurance plan. Copayment is a capped personal contribution when medical services are accessed (sometimes referred to as a top-up fee).

User fees are similar to copayments. The term is often used to refer to OOPs payments for public health care. User fees or direct out-of-pocket payments have been introduced or extended in many developing countries as an alternative form of health care financing. (See World Bank 1993 and ILO (2000, Chapter 4)). According to Mills (2007), direct payments are 'an extremely important source of health financing in many countries. Usually they constitute direct payment to providers independent of other sources of funding, though in some countries sizeable co-payments may be required by SHI [...]' WHO (2005) provides evidence and argument for avoiding direct payments on equity grounds.

James (2006)) distinguishes between user fees ('official fees charged by public health providers for basic as well as higher level services as used in one form or another in most countries and contexts) and out-of-pocket (OOP) payments, which may include charges by private, NGO and community-managed services.

Informal or unofficial payments are payments that do not have an official sanction. Payments can be made for medical supplies, accommodation or to augment clinical salaries. They are associated with, but not limited to, health systems where major public budget cuts have been made (e.g. the transition economies).

User fees can involve targeting services on the poor by the use of waivers or exemptions from charging. User fees can involve targeting services on the poor by the use of waivers or targeted charges. The World Bank has argued the need for a “pro-poor” policy stance involving targeted charges: “A comprehensive approach to poverty reduction...calls for a program of well-targeted transfers and safety nets as an essential complement to the basic strategy” (World Bank 1990:3).

Analysis of user charges

User charges are the most regressive financing method because they do not allow for pooling. In the case of user charges the risk pool is an individual or a family unit. Charges are justified as we have seen as an alternative to scarce state funding and as a means of controlling excess demand (‘moral hazard’). There is a large literature on their effects on access to care and their inefficiency. Yet charges remain the main method of funding health care in low and middle income countries. In eastern Europe, for example, user charges have been increased substantially in the last 10-15 years. According to Figueras et al (2008: 40), in 2004, ‘out-of-pocket spending constituted over half of total health spending in Armenia, Azerbaijan, Georgia, Kyrgyzstan, Tajikistan and Uzbekistan (176). More recently, cost sharing has been extended elsewhere, including in several EU Member States (Austria, the Czech Republic, Estonia, France, Germany, Hungary, the Netherlands, Portugal and Romania). In Africa, the 1987 Bamako Initiative introduced user fees.

User fees have been widely criticized as inequitable and counter-productive (Creese, 1997). The Bamako Initiative adopted by several African countries in 1987 provides an example of the mixed results of user fees. The initiative was intended to improve the utilization and quality of primary care but subsequent research showed:

- the revenue generated fell well short of expectations
- exemption policies for the poor were inefficient (in Ghana, for example, ‘less than 1 in 1000 surveyed was granted exemption when an estimated 15-30% lived in poverty’ (Singh A (2003) Building on the user fee experience: the African case. Geneva: WHO)
- revenue in most cases did not remain at the community level for which it was intended
- one third of countries reported a fall in primary care utilization and one third reported an increase (Elgazzar, 2007).

There is a strong correlation between high OOP payments and the percentage of families facing catastrophic health expenditure (defined as 40% or above of non-subsistence income ((income available after basic needs, such as food, have been met).

Waiver systems, which exempt the poor or poorest from charges, have failed to overcome problems of access and an alternative approach is to provide public financing only for services aimed at the poor. Service level as opposed to finance targeting has become an important component of external financing of health

care in developing countries. This change of approach complements and is reflected in the PRSP process (Bitran, 2003).

Moreover, not all user charges allow for waivers. For example, unofficial payments are by definition outside a system of controlled charging. Ensor (2004: 239) deals with the problem of informal or unofficial payments in transition country health systems. Seen as a transitory problem in the early stages of the reform of communism, by 2004 they were attracting more attention from policy makers. The prevalence of payments is difficult to assess but the paper cites examples of patients unofficially contributing towards the cost of drugs, food and other supplies, and also to salaries:

‘In Albania, reports suggest that all cadres of hospital and clinic staff receive payments, the highest going to those carrying out specialist procedures such as heart surgery. In Kyrgyzstan, the same survey reported earlier also found that 25 per cent said they made a ‘gift’ to staff. In Bulgaria, Delcheva, Balabanova, and McKee (1997) suggested the unofficial cost of an operation accounted for more than 80 per cent of the average monthly wage. Further a-field, in Uganda, it is suggested that unofficial charges can double a health worker’s wage. A study in Poland suggested that the majority of unofficial payments are paid directly to physicians for services although it is possible that some are shared between other staff.’

Payments of this type are often seen as responses to the budget reductions following reform. However, Ensor suggests that there may be other explanations such as custom and practice.

Charges can also be analyzed for their effects on rational drug use. A recent Cochrane review concludes: ‘Introducing direct co-payments reduced drug use across studies. Patients responded by discontinuing drugs or by paying an increased proportion of the costs themselves. Reductions were found for life-sustaining drugs or drugs that are important in treating chronic conditions as well as other drugs, suggesting that patients may not have been able to prioritise their drug use when faced with a reimbursement restriction.’ (Cochrane Collaboration (2009)) The review also suggests that charges are counter-productive if its adverse effects on health lead to a greater call on health services.

