Possible objectives and resulting entitlements of essential health care packages

Neil Söderlund *

Centre for Health Policy, University of the Witwatersrand, P.O. Box 1038, Johannesburg 2000, South Africa

Received 20 March 1998; accepted 27 July 1998

Abstract

The notion of a defined ‘core package of essential health care services’ has appeared in many different health reform proposals in the 1990s. This paper attempts to explore the possible objectives of the ‘core package’ component of health care reform. Two broad applications are apparent: the use of essential packages to ration scarce public funds and the incorporation of a minimum benefit package into ‘managed competition’ type reforms, where they constitute a mandated minimum level of private insurance cover. Eight possible objectives for an essential benefit package are described: To protect against catastrophic illness events; to ensure social risk pooling; to improve allocative efficiency in the health system; to eliminate ‘high burden of disease’ conditions; to improve equity of access to services; to combat cost-escalation; to encourage competition between insurers; and to facilitate public participation and transparency in decision making. Closer examination of objectives reveals that they often conflict, which suggests that a clear understanding of the purpose of reform is essential before it is worthwhile devoting energy to the development of essential benefit packages. It is argued that two main clusters of objectives emerge from the eight described, representing Rawlsian (risk avoidance) and utilitarian (efficiency improvement) social welfare philosophies, respectively. Practical experience suggests that priority setting exercises have been unsuccessful in meeting efficiency objectives, but that they may well be quite useful in fulfilling risk-pooling aims. © 1998 Elsevier Science Ireland Ltd. All rights reserved.

Keywords: Essential health-care package; Cost-effectiveness; Priority setting; Rationing; Risk-pooling

* Tel.: +27 11 4899883; fax: +27 11 4899900; e-mail: soderlun@icon.co.za

0168-8510/98 - see front matter © 1998 Elsevier Science Ireland Ltd. All rights reserved.

PII S0168-8510(98)00039-6
1. Introduction

The 1993 World Development Report [1] has generated a significant interest, especially in the developing world in the concept of essential benefit packages as tools for rationing publicly funded health care. This notion of explicit packages for rationing purposes is, of course, nothing new and forms an integral part of debates that have existed in tax-funded health care systems for some time. In these systems, the debate often vacillates between macro rationing (i.e. a universal package of priority services for all) versus micro rationing, where guidelines, protocols and prior-authorisation by third party payers determine which individuals have access to which services in which disease context, usually on a case by case basis.

More recently, the essential package concept has been reformulated as a regulatory component of managed competition initiatives, such as those advocated by NERA [2], Enthoven and Kronick [3] and applied within the Dutch Health Insurance Reforms of the early 1990s [4,5]. That is, government mandates a minimum acceptable level of health care cover which all, or a subset of citizens, are required to take out from one of a number of vendors. Those who cannot afford the minimum level of cover are either excluded from the mandate, or their access to cover is subsidised by the state. The regulated minimum package of cover is determined by some notion of health care priorities, but these might not be the same as those used to ration taxpayers money.

A review of the literature reveals that the ‘core package’ concept has been touted as a near universal fix-it for a multitude of health system ailments [2,6,7]. There is also the assumption that, fickle public opinion aside, most core packages would end up looking fairly similar. Physicians, of whom this author can claim to be a distant relative, have an inbred suspicion of universal cures and this was probably the main factor prodding deeper investigation. I have managed to identify eight potential objectives for use of the essential benefit package both as a rationing and a regulatory tool. Many of these objectives are not mutually compatible and far from being a cure-all, the core package approach appears to require specific policy goals before it can be of much use at all. The section which follows outlines the range of potential objectives of the minimum package approach in both its ‘rationing’ and ‘regulatory’ guises. Following this, I have tried to examine the degree of compatibility between different objectives. The concluding section identifies sets of more or less coherent objectives for health systems with similar needs.

2. Possible objectives of minimum benefit package regulation

2.1. Facilitating catastrophic insurance cover

For conditions where costs of treatment are high, but isolated and relatively rare, very few households could reasonably expect to meet these out of current income or savings. In these situations, households might have to forego other expenditures to their long term disadvantage (i.e. on food supplies or education) in order to
cover costs [8,9]. In many cases, however, these costs would be quite affordable if shared, through insurance arrangements, between individuals and over time. As Hammer and Berman point out, market failure is common in this context for a number of reasons [10]. While a natural market might well develop for insurance covering lower cost, higher frequency health events, the need for large risk pools and a lack of knowledge regarding real probability of catastrophic events often inhibits the spontaneous development of such cover (assuming that it is affordable). Furthermore, adverse selection is another source of market failure, resulting in those who perceive themselves to be at low risk taking out little or no cover. In these situations, governments might intervene to mandate insurance for catastrophic (high cost, but low frequency) health events. The case of catastrophic costs impacting on household viability is perhaps of most relevance to very poor countries, where no state welfare mechanisms exist [8].

This problem is further complicated by government initiatives to provide free care to the poor, often by means of public hospitals. In addition to providing care to the poor, the state often becomes the de facto insurer of last resort for catastrophic events for the whole population. If it willingly takes on such a role, this is not a public policy problem. In many cases, South Africa being an example, the state plays this ‘last resort’ insurer role with some reluctance. In South Africa, Government has explicitly stated that the role of free state hospitals is to provide for the poor and has a stated policy of charging the voluntarily uninsured user fees. Since public hospitals cannot turn away patients in need and have difficulty applying and collecting means tested user fees, a free-rider problem develops. This, in turn generates a demand for private insurance which at first appears counter-intuitive, with consumers rejecting cover for large risks and buying cover for predictable, low-cost, minor ailments instead. Fig. 1 shows two possible insurance

Fig. 1. Minimum insurance objective for an essential benefit package.
policies covering different types of illness event, classified in terms of their total
costs and likelihood of occurrence. The state cannot refuse treatment for conditions
in area A and consequently, the private insurance market covers conditions in area
B first and only extends into area A (i.e. substitutes for state services) when levels
of cover expand. The low-cost medical schemes sector in South Africa illustrate this
point graphically. They offer comprehensive cover for general practitioner, phar-
macy, dental and optometry benefits (none of which are routinely provided by the
state) and minimal cover for high cost hospital treatments, which public hospitals
cannot refuse to provide [11,12].

Under the scenarios described above, a possible objective for a minimum benefit
package would be to ensure that catastrophic events are fully covered first, before
cover for low-cost or predictable types of care can be made available. In other
words, if it is stipulated that insurance has to offer, as a minimum, cover for high
cost, non-discretionary emergency care, then household calamity and/or free-rider
problems largely disappear. In practice, meeting this objective involves providing
first for conditions that are severe and/or require urgent treatment and/or for which
treatment is very costly.

2.2. Ensuring risk-based cross-subsidies

Many countries have included within health care financing regulation attempts to
ensure that all citizens can purchase affordable basic insurance cover regardless of
their health status or past claims experience. Two related phenomena serve to
fragment health insurance risk pools and eliminate the cross-subsidy from healthy
to sick enrollees. Adverse selection refers to the voluntary opting out of insurance
cover by low-risk enrollees. Underwriting and risk selection, on the other hand, refer
to the actions taken by insurers to minimise the degree of cross-subsidy required of
low-risk enrollees and so make insurance cover more attractive to them. The well
known Rothschild–Stiglitz model [13], of risk-pooling in insurance suggests that
vendors will sell more insurance the closer their premiums approximate actuarially
fair prices (i.e. they reflect ex ante risk of illness).

Mandatory cover is one commonly used approach to reducing adverse selection.
So called ‘community -rating’ of premiums, open enrolment laws, risk-equalisation
funds and high risk pools are all mechanisms designed to further effect cross-subsid-
ies from low- to high-risk members. In exchange for their current support of the
elderly and sickly, young and healthy contributors are implicitly guaranteed that
when they become old, they in turn will be subsidised by a future generation of
young members.

Regulation of premiums alone is insufficient to ensure risk pooling, since by
designing policies that appeal specifically to different risk groups (so-called ‘separat-
ing contracts’), insurers can effectively distinguish high- from low-risk members and
place them into different risk pools with differentiated premiums. They might do
this, for example, by selling two policies, only one of which offered comprehensive
cover for long-term nursing care. Since young people (low-risk) think themselves
unlikely to require these benefits, they would choose the policy without nursing care
benefits, whereas the elderly (high-risk) would likely purchase nursing home cover. Assuming effective separation, ‘actuarially fair’ premiums can be charged to each for all benefits received (not just long-term nursing care) and the extent of cross-subsidisation minimised.

Appropriate regulation of benefits is thus an essential accompaniment to premium regulation. If the conditions that predominantly affect high-risk groups are included in the core benefit package, then low-risk groups are forced to cross-subsidise high risks for these conditions and risk-pooled premiums are guaranteed at least for the core package. A package that has these objectives would include primarily, conditions whose incidence is concentrated in easily identifiable high-risk groups, rather than spread evenly over the population. This approach to package development may be illustrated using AIDS and trauma treatment. Insurers might identify future AIDS sufferers by offering two treatment packages, with and without treatment for opportunistic infections. Enrolees who know themselves to be at risk of AIDS will tend to choose the former, while those who are at low, or zero risk, would choose the latter. By putting treatment of AIDS-related disease into the core package, therefore, one removes choice in taking out this area of cover and forces risk pooling. On the other hand, apart from certain high risk occupations, there are few ways of predicting whether one is likely to suffer serious trauma. Individuals are thus likely to be relatively willing to pool risks for trauma and this benefit might be left out of a minimum package, where the prime aim was to combat risk fragmentation.

2.3. Improving allocative efficiency

Economists conventionally define production as allocatively efficient when factors are combined in their optimal proportions. If we treat attributable health status improvement as the output of health care, then the factors of production may be viewed as the different modalities and types of treatment that can be administered to different disease states.

This is illustrated in Fig. 2, which describes a simple world with three diseases, for which effectiveness of treatment differs. Health gain from treatment is plotted on the vertical axis versus the amount of treatment administered on the horizontal axis. The solid lines A, B and C represent expected health gain from different levels of treatment input for the three diseases. If we have three individuals, suffering from A, B, C, respectively—and assuming a utilitarian social value function—then the optimal allocation rule between spending on these diseases, assuming a fixed budget, would be to equalise the marginal productivity of the last dollar spent on each condition. This marginal productivity is represented by the slope $S$.

Under optimal conditions, disease B gets more spent on it ($X'$) than disease A ($X$) and C attracts no health care spending since $\max[\delta Y_C/\delta X] < S$ for all positive values.

Dealing only with A and B, total health gain is:

$$Y_1 + Y_1'$$
If, on the other hand, a simple cost based limit is applied, under the same resource constraints, $1/2(X + X')$ is spent on each patient, yielding a total health gain of:

$$Y_2 + Y_2'$$

or

$$(Y_1 + Y_1' + p - q)$$

To demonstrate the that $Y_1 + Y_1'$ is a higher level of output, therefore, we simply have to prove that $q > p$. At $1/2(X + X')$:

$$\frac{\delta Y_A}{\delta X} < S$$

and

$$\frac{\delta Y_B}{\delta X} > S$$

dependent

$$\delta Y_A < \delta Y_B$$, or $q > p$

By stipulating a higher level of cover for disease B than A and no treatment for disease C, the minimum benefit package thus has the potential to enhance the allocative efficiency of health care spending. This notion of maximising national health care productivity is the central justification given by the 1993 World Development Report [1] for recommending a core package approach to public health care spending. This rationale has been extended into the development of essential care packages for a number of developing countries including Mexico,
Colombia, Uganda, China and others [14,15]. Obviously, for many conditions and treatments, marginal benefit curves are not known and there remains individual variability in treatment response. In extreme cases, such information is known with relative certainty and this may allow exclusion of particularly bad buys, or inclusion of good buys, if not the complete ranking of all possible interventions. The development of packages of entitlements based wholly or mainly on cost-effectiveness has yet to be seen at a national level anywhere in the world.

2.4. Reducing burden of disease

The 1993 World Development Report uses two important criteria for inclusion of interventions in its package of essential benefits: cost effectiveness and burden of disease. The former is clearly justifiable and present in virtually all attempts to construct basic packages of services [7], although difficult to apply practically. It is not immediately obvious why the latter is not redundant in the presence of perfectly efficient care. If the number of disability adjusted life years gained per dollar spent is the same for two different interventions, then this objective would suggest the commoner of the two should receive higher priority. Rare conditions, even if cost-effectively treated, would thus be left out of the package. It would seem appropriate to use ‘total burden of disease reduced’ to inform package definition in two circumstances only:

1. Where there are significant economies of scale associated with an intervention. These are likely to occur where significant fixed costs have to be met up front in order to deliver the service concerned. Examples might include the use of lithotripsy machines to treat kidney stones or the introduction of an immunisation schedule which requires retraining of primary care nurses. Health interventions, validly thought of as ‘programmes’ rather than treatments to individuals, tend to have these high start-up costs. The scale economies argument clearly applies also to new high-tech interventions. In both cases, the intervention might be hard to justify unless start-up costs can be distributed across many potential beneficiaries, or a large ‘aggregate burden of disease’. For most health care interventions, the costs per patient of treating rare, as opposed to common conditions, do not differ much.

2. Where there are significant positive externalities (either interpersonal or intertemporal). Where a service yields positive externalities, consumers are likely to underinvest in cover, since their personal benefit will be less than the cost associated with optimal social levels of provision. Government intervention in such areas is justified to optimise the level of provision of such services. For example, individual willingness to pay for tuberculosis treatment will only take into account the health gains to the sufferer and not those accruing to persons who would otherwise have been infected by the person concerned. Prioritising between interventions, for which strong positive externalities exist, needs to take into account community ‘burden of disease’, rather than individual self-perceived need for care.

Where these circumstances do not exist, simple comparison of cost-effectiveness of interventions would appear to achieve the same end.
2.5. Improving equity

Elimination of differential access to health services is a commonly voiced objective of reforms. Most commonly, inequity in access is income related. Core packages with rationing objectives (such as the Oregon initiative) have often sought to equalise access by shrinking the total size of the package and hence increasing the number of covered persons. Even within the regulatory context, however, minimum benefits might have a role to play in enforcing social solidarity. South Africa, with its history of racial and geographic inequities, provides the potential for such effects. Despite appropriate resource shifts towards needy groups, institutions and their staff might be tempted to leave the status quo of inferior access or quality of care for less demanding patients unaddressed. Many health insurers in South Africa, for example, have excluded from cover high cost illnesses affecting mainly black members (e.g. AIDS and tuberculosis), but not those affecting white members (e.g. coronary artery disease) [16]. Lower levels of expectation in disadvantaged groups enhance inertia in the move to equity. In this context, a set of minimum benefits becomes a powerful tool in enforcing equity benchmarks. Evidence that some patients do not get access to essential care, whereas others are provided with non-essential treatment, is difficult for state, insurers and providers to refute.

2.6. Controlling moral hazard and cost escalation

Cost escalation in excess of consumer price inflation and often unexplained by demographic or morbidity changes, is evident in many industrialised and middle income countries. In the years 1990–1995, health care inflation in South Africa averaged 4% above the consumer price index for the insured sector [12]. Control of cost-escalation has undoubtedly been a major stimulus behind essential benefit package initiatives in developed countries [17–19]. Three arguments are used to link benefit limits to cost control. The first and most commonly argued does not apply to the regulatory context discussed here, while the second and third may well be of value. Firstly, it is argued, the simple act of limiting benefits, regardless of how it is done, must bring down costs (or at least costs incurred by one particular source, generally the state). Objectives underlying cost control may include expanding health service coverage to a wider range of people (as was the case in the Oregon Medicaid experiment [20]). This crude approach to reducing state health care spending may not reduce overall health care costs, if costs are displaced from a relatively efficient financing or provision mechanism to an inefficient one [21]. Furthermore, in both the rationing and regulatory contexts, the intent is to design a minimum package, with people being free to purchase more comprehensive cover if they wish, so this cost-control mechanism is unlikely to apply. With a regulated minimum package, there is no reason to believe that someone who already has more extensive cover than the core package would be affected at all by the minimum benefits regulatory initiative.
The second and probably more important mechanism by which the core package approach may control costs is via the exclusion of more discretionary types of care. Certain types of health care appear particularly prone to moral hazard, both supplier and consumer induced. We might regard such health care as largely ‘discretionary’ in nature and empirically, such interventions often manifest with wide variations in utilisation rates between different practitioners, geographic regions and financing environments, without obvious differences in need [22–24]. Examples include hysterectomies for post-menopausal bleeding, surgery for backache and tonsillectomies in children. In the South African context, public and private sector providers show similar age–sex standardised utilisation rates for appendectomies, hospitalisation for pneumonia and treatment of fractures. Tonsillectomy rates in populations insured for private sector care are more than 20 times those of the populations served exclusively by the public sector [25]. Defining an essential package consisting mainly of non-discretionary services is one approach to limiting moral hazard and its consequent impact on cost-escalation.

The third mechanism whereby minimum package regulation may reduces costs is via risk pooling objectives mentioned in Section 2.2 above. If regulation mandates risk pooling for a set of defined benefits, then insurance theory suggests that, on average, less comprehensive cover will be chosen by enrollees [13,26]. It is well recognised that less than full insurance coverage is desirable to offset moral hazard problems [27,28]. Consequently, by facilitating a single, community-rated risk pool, the core benefit package might also, indirectly, reduce moral hazard and cost-escalation.

2.7. Fostering competition

By mandating a level of cover, rather than membership of a particular fund, or indeed, contribution to the tax system, the core package approach stimulates competition between different insurers. The South African health insurance market shows an enormous amount of product differentiation for a relatively small potential enrollee population. Over half of the 170 or so health insurance mutual funds have <5000 members and few, if any schemes offer exactly the same benefit package. Enthoven describes a similar situation in the US market [29]. In such situations, price comparisons are virtually impossible. With regulation to ensure a minimum benefit package, it is likely that many, if not most schemes will offer the essential package as one of their products and at least at this level, price comparisons may have some value. Of course, if a set of essential benefits is specified together with a mandate on coverage, as is the case in South Africa [30], then the minimum benefit package is an essential prerequisite for fair competition between insurers. Without minimum benefits being specified, insurers would attempt to sell increasingly stripped down packages (at lower costs) to employers and employees wishing to effectively avoid the mandate.
2.8. Facilitating transparency and participatory democracy

Public participation has formed an essential component of package determination where the intention is to ration public health care resources. In a rationing context, the public participation and transparency objectives are difficult to dispute. If societies are voting for and/or paying for health care (amongst other) benefits, they deserve an explicit description of what is being provided and a voice in shaping this [31]. Much of implicit rationing in the past was hidden behind the veil of clinical decision making and aided by widespread public ignorance. Public understanding and articulation are rapidly increasing as is professional reluctance to take on the role of rater. When public funds are being used, there is a case for explicit definition of benefits even in the absence of the abovementioned efficiency, equity and cost-control gains. Put another way, even if there is no explicable logic to the construction of an essential package, the act of specifying benefits might constitute an improvement in societal well-being.

It is less clear whether the facilitation of transparency and public participation could be claimed as an objective for the regulatory form of the minimum benefit package. In this context, individuals (and/or their employers) are paying the cost of the package, not the state and it would be in their interests to shrink the package to its smallest possible size in order to maximise their freedom to choose their health care cover package. Furthermore, it is unclear what participation those not covered by the minimum benefit mandate (e.g. the unemployed) should be entitled to. Benefits in terms of social cohesiveness and a more widespread public awareness of necessary choices in health care spending might, however, be significant.

3. Compatibility of different core package objectives

Clearly, no single package of benefits can meet all of these objectives. Some might be more compatible than others. The potential conflict between efficiency and insurance objectives has already been well demonstrated by Hammer and Berman [10]. When comparing two interventions of equal effectiveness, efficiency objectives would dictate the selection of the lower cost option, whereas minimum insurance cover objectives would dictate the selection of the higher cost option, since this is least likely to be affordable out-of-pocket. Similar comparisons can be made between other objectives. Public participation objectives might conflict with many other objectives. For example, a public majority might support excluding an intervention which affects a small identifiable minority of members, even if treatment meets cost-effectiveness and severity criteria. Other prioritisation exercises have shown such biases with regard to treatments for alcoholic liver disease [32] and chronic psychiatric care [33] and many mutual health insurers in South Africa, (whose benefits are determined by boards representing members), have issued blanket exclusions on HIV related care, possibly for the same reasons. When the purpose of the package is simply to regulate the type of cover individuals can buy on the open market (rather than actually contribute towards it), public preference
would probably be for less, rather than more regulation and hence a smaller basic package.

The extent to which any one objective dominates package definition will reflect what the role of the package is seen to be. For example, if the prime objective is to prevent dumping on state services, then degree of emergency and unaffordability of treatment might be important criteria [10,34]. If, on the other hand, improving health service efficiency is of prime concern, then effectiveness, cost effectiveness and burden of disease should define the package [1].

Olsen has usefully attempted to classify priority setting initiatives according to underlying beliefs about social justice [35]. Cost effectiveness, burden of disease reduction and moral hazard elimination objectives fit squarely within a utilitarian normative framework. Objectives around risk pooling, equitable access and catastrophic cover, on the other hand, would appear to reflect a preference for a Rawlsian distribution of social goods, wherein most efforts are directed at improving the condition of the worst off. The (non-altruistic) justification for Rawlsian social preferences is that all individuals, fearing that they might end up in the position of the worst off, seek to improve this position ex ante. That is, preference for a Rawlsian distribution simply reflects risk-aversion¹. The trade-off between efficiency and risk avoidance (or utilitarian and Rawlsian social distribution preferences) seems to be a useful way of classifying most of the eight specific objectives for essential health care benefit packages. This natural conflict has been well recognised from the earliest prioritisation exercises [36].

Most initial applications of the essential package philosophy were aimed at informing the rationing of scarce public funds [37–41]. They set out with strong efficiency improvement objectives, no doubt influenced by the boom in micro-economic evaluation of health care interventions over the last 10–15 years, but often ended up, as a result of public, professional and political pressure, incorporating risk avoidance aims. More recently, governments have recognised the a priori validity of risk avoidance objectives. Examples include Norway [35] and Sweden [42]. Where regulation aims to eliminate free-riding on public health services via an insurance mandate for a minimum package, risk aversion objectives would seem to be particularly important. In developing countries, however, largely following the initiative of the 1993 World Development Report, priority setting exercises have emphasised efficiency objectives to the virtual exclusion of risk-aversion ones [14,15].

In addition to the theoretical argument favouring risk pooling, rather than efficiency objectives, practical considerations suggest that ‘efficiency’ packages are less than ideal tools. Empirical evidence derived from applying efficiency and risk avoidance criteria to existing hospital care in South Africa [25] has shown that the main difference between the two packages is that elective surgical interventions such as hip replacements and hernia repairs selected in the former are displaced by costly

¹ Olsen suggests that true egalitarian social welfare preferences have little relevance to health priority setting because of their insistence on equal distributions of outcomes, which, in the case of health, is obviously impossible.
cancer therapy, neonatal intensive care, renal dialysis etc., under the latter. The cost savings derived from excluding elective surgery are far greater than the cost increases due to high cost interventions given current practice patterns, implying that a ‘risk aversion’ package would provide greater immediate cost savings than an ‘efficiency’ package. One possible reason for this observation is that very little inefficient care exists. High levels of cost escalation without much increase in health status suggest that this is unlikely. A second and more convincing explanation is that inefficient care is not well characterised and eliminated by crude categorisation and ranking type exercises. Within any category of care, there is likely to be a wide range of efficiencies of the same treatment depending on a multitude of micro-clinical considerations. This fact has formed the mainstay of criticisms of the Oregon and other similar rationing exercises [43,44]. Guidelines, protocols and appropriate incentives for clinicians are probably much more appropriate tools for efficiency improvement than essential benefit packages. Discrete categories can, however, quite accurately capture the degree of severity, urgency and discretion associated with a given intervention for a specific condition. While caesarean section might vary in its cost effectiveness as a means of delivery, its implied level of urgency does not. Optical cataracts are never life threatening, or irreversibly disabling, but their removal still varies widely in cost effectiveness depending on the level of disability caused.

While essential benefit packages started out as an attempt to improve health service efficiency, theoretical and practical factors suggest that they are actually more suited to preventing catastrophic illness events, ensuring risk pooling and eliminating inequitable access to care, especially where they form part of regulatory, rather than rationing exercises. This appears to have been recognised by developed country initiatives, but many developing countries are still attempting to improve public and private health system efficiency by means of the core package approach.

Acknowledgements

This work was funded by the Anglo–American Corporation of South Africa and the Partnerships for Health Reform (PHR) Small Applied Research Grant Program, with funding from the US Agency for International Development (USAID). Thanks to Alex van den Heever, Brian Brink, Roland Bright and participants at the 1997 Economics Society of South Africa conference for useful insights and comments. Any errors or omissions remain the responsibility of the author.

References


