Health technology assessment (HTA) and economic evaluation: efficiency or fairness first

Jeff Richardson & Michael Schlander

Published online: 20 Dec 2018.

Submit your article to this journal

Article views: 1005

View related articles

View Crossmark data

Citing articles: 2 View citing articles

Routledge
Taylor & Francis Group

ARTICLE

OPEN ACCESS · Check for updates

Health technology assessment (HTA) and economic evaluation: efficiency or fairness first

Jeff Richardson [a] and Michael Schlander [b,c]

aCentre for Health Economics, Monash University, Melbourne, Australia; bDivision of Health Economics, German Cancer Research Center (DKFZ) & University of Heidelberg, Heidelberg, Germany; cInstitute for Innovation and Valuation in Health Care, Wiesbaden, Germany

ABSTRACT

The economic evaluation which supports Health Technology Assessment (HTA) should inform policy makers of the value to society conferred by a given allocation of resources. However, neither the theory nor practise of economic evaluation satisfactorily reflect social values. Both are primarily concerned with efficiency, commonly conceptualised as the maximisation of utility or quality adjusted life years (QALYs). The focus is upon the service and the benefits obtained from it. This has resulted in an evaluation methodology which discriminates against groups and treatments which the population would like to prioritise. This includes high cost treatments for patients with rare diseases. In contrast with prevailing methods, there is increasing evidence that the public would prefer a fairness-focused framework in which the service was removed from centre stage and replaced by the patient. However methods for achieving fairness are ad hoc and under-developed.

The article initially reviews the theory of economic evaluation and argues that its focus upon individual utility and efficiency as defined by the theory omits potentially important social values. Some empirical evidence relating to population values is presented and four studies by the first author are reviewed. These indicate that when people adopt the social perspective of a citizen they have a preference for sharing the health budget in a way which does not exclude patients who require services that are not cost effective, such as orphan medicinal products (OMP's) and treatments for patients with ultra-rare diseases (URD's).

Introduction

In the context of Health Technology Assessment (HTA) the theory of economic evaluation seeks to show how best to allocate the health budget to maximise individual and social welfare. It is argued here that these objectives are not achieved. In the first part of the paper it is shown that there are serious defects in the theory. The result of these is that the evaluation methodologies derived from the theory discriminate against some patients and services. In particular these include patients with rare diseases who require high cost services which are not found to be cost effective using these methodologies. In the second part of the paper the results of some empirical studies are reviewed. These support the belief that the efficiency focused paradigm of economic theory does not reflect the fairness focus of population values.

The assessment of individual and social welfare require criteria. and economics has adopted the theory of preference utilitarianism, that individuals seek to maximise their utility which is defined by the strength of their preferences for different options. Economic Welfare Theory – the theoretical edifice from which evaluation theory is derived – extends this assumption to the doctrine that social welfare is a function only of individual utilities. This is commonly simplified to the utilitarian objective of maximising (unweighted) utilities.

The methods used to support HTA require measurable goals. This has resulted in three main types of analysis, each derived from Welfare Theory. Cost benefit analysis (CBA) seeks to measure different outcomes using people's willingness to pay for them on the assumption that the willingness to pay measures the strength of a person's preferences. As it also reflects a person's ability to pay, the preferred forms of analyses in the context of a communal health system are cost effectiveness analysis (CEA) – when information relating to the quality of life (QoL) is unavailable – or, increasingly, cost utility analysis (CUA). In this, the quality of life (QoL) of different health states is measured by the utility – the strength of preferences – for different health states. Consistent with the original formulation

of utilitarianism by Jeremy Bentham the maximum utility which an individual may achieve each year in CUA is set equal to 1.0 and the unit of benefit is the quality adjusted life year (QALY) which is defined as life years multiplied by the utility of the life years. QALYs may therefore be conceptualised as the total utility obtained during those life years. The objective of CUA is then to maximise the number of QALYs – utilities – gained [1]. This is achieved by selecting services which minimise the cost per QALY. The utility of different health states is generally measured with one of the available multi attribute utility instruments (MAUIs), the most commonly used being the EQ-5D [2].

Despite its analytical elegance Welfare Theory encounters a number of conceptual problems. The most succinct and comprehensive summary of these is given in the first edition of the text by Tom Rice [3]. Three problems of particular relevance for the present article are outlined below. The common theme is that current theory and practise exclude a number of considerations which would extend the range of services which the public might wish to include in a communal health service.

The first problem arises from the assumption that an individual's wellbeing is maximised by maximising utility. However the concept and measurement of utility and the evidence that its maximisation also maximises wellbeing are all problematic. 'Utility' has been used to refer to different concepts. Originally Jeremy Bentham [4] and Francis Edgeworth [5] defined it in terms of happiness, the tendency to increase pleasure or reduce pain (hedonic utilitarianism) [6]. These concepts are measurable. Subsequently, however, the concept morphed into the strength of people's preferences (preference utilitarianism) and in the context of uncertainty, preferences which are consistent with the Expected Utility Hypothesis, discussed below [7,8]. More recently a distinction has been drawn between 'decision' and 'experienced' utility, preferences before an outcome is known and wellbeing after it has occurred [9].

Welfare Theory assumes that maximising utility – defined as decision utility – also maximises individual wellbeing. But the assumption cannot be directly tested as decision utility cannot be observed. The problem is overcome in Welfare Theory by invoking the 'revealed preferences' doctrine: that people's preferences are revealed by their choices. As individuals are assumed to be best placed to assess their own wellbeing, the choices which maximise utility also maximise wellbeing. However without further evidence the solution is a tautology. Choices are made. The motivation for them is that they maximise wellbeing. But the evidence that they maximise wellbeing is that the choices are made. Either half of the

tautology appears plausible: choices are motivated by the attempt to maximise wellbeing; the evidence that wellbeing is maximised is that choices are made. Together the two halves are vacuous: choices are made because choices are made. However choices are the result of numerous influences which may or may not lead to maximum wellbeing – habit, duty, religion, peer pressure, marketing, misinformation, the inability to assess the consequences of choices, etc. In the context of choices relating to health the last two factors are particularly important.

Nevertheless, while the relationship between choice and wellbeing may be unclear, maximising 'utility' can be defended by the ethical judgement that we should respect choice. But this is clearly unacceptable in cases where choice is foolish, misinformed or harmful, possibly as a result of an individual's mental instability. This implies that understanding the determinants of choice is important for the argument that it should be respected, and that utility should be maximised.

Recently behavioural economics has investigated the motivation for choice. Relevant findings include the results from the ultimatum and dictator games [10–15]. These indicate a powerful propensity to share benefits and this will be reflected in people's social preferences; how they would like social enterprises, such as a collective health scheme, to be managed. However Welfare Theory is also based upon individualism. Utilities are a function of the goods and services consumed by the individual. They are not a function of other people's situations. More particularly there is nothing in the methods of CUA which can take account of socially induced motivations such as the ubiquitous propensity to share discussed later in the paper. Utility measured by MAUI's are derived from choices which only take account of the quality and length of life.

Personal preferences are also influenced by the uncertainty of future events. The orthodox response to this is to assume that people's choices under risk are explained by the expected utility hypothesis, that people maximise the (mathematical) expectation of utility (the sum of the probability weighted utilities which might eventuate). All else equal this re-establishes the maximisation of utility or QALYs as a personal goal. The utility/QALY maximising mix of services will also maximise expected utility/QALYs. Conversely, if the utility maximising mix is what people choose then the expected utility theory must be true [16]. However it has been known for over 60 years that the theory does not describe individual preferences when there is uncertainty [17]. Consequently, evaluation theory does not properly address an important element of what an insurance scheme is intended to achieve, namely

people's aversion to risk and uncertainty and the possibility of experiencing a severe health state which requires a high cost treatment.

The importance of uncertainty and socially induced motivations are not directly proportional to the number of people receiving a service. This is also true for other potential benefits such as compassionate externalities and the option value of a service's availability. Nevertheless decisions in CUA are based upon the incremental cost effectiveness ratio (ICER) – the incremental cost/QALY. In calculating this the number of patients enter both the numerator and the denominator and cancel out. Consequently, the possibility that high cost services for rare diseases may be valued for benefits which do not depend upon patient numbers is ruled out by the construction of the decision criterion. The cost/QALY ratio cannot vary with the rarity of the disease.

The contestable assumptions of Welfare Theory – or 'welfarism' – call into question the theoretical foundations of economic evaluation which it provides. This has led to the creation of an alternative theoretical rationale for evaluation methods which has been described as 'extra-welfarism'. This changes the goal posts from the maximisation of utility to the maximisation of health. QALYs – utility times life years – remain the unit of measurement but the focus shifts from utility to the QALY which is assumed (controversially) to be a measure of health per se.

While welfarism sought to measure the fundamental goals of economic activity, extra-welfarism only claims to describe the objectives of the health sector. While Culyer [1,18] has drawn upon Sen's capability approach to wellbeing in support of extra-welfarism the case for the *maximisation* of health (as distinct from its importance) and for disregarding other considerations is only based upon the rhetorical assertion that the purpose of the health sector is the maximisation of health. The assertion is not backed by theoretical argument or empirical evidence of population support for QALY maximisation [19,20].

The second problem occurs with both welfarism and extra-welfarism. As satirised by Uwe Reinhardt, maximising efficiency disregards fairness and the distribution of benefits [21]. Economic evaluation results in winners and losers. Winners are those who require services which are cost effective. Losers are those requiring services which are not cost effective. To maximise QALYs, resources must be provided to the first group but not to the second. However a rational individual who evaluates alternatives from behind a 'veil of ignorance' (ie not knowing what he or she may need in the future) might select insurance which, for example,

provided 4 people with 3 QALYs rather than insurance providing 3 of the 4 people with 5 QALYs and the fourth person with nothing, despite the latter option maximising the QALY gain. From a self-interested perspective the individual might be the fourth person in the future.

Welfare Theory has approached this problem in two ways, both of which are unsatisfactory. The first approach has been to postulate that there is a 'social welfare function' (SWF) which incorporates distributional preferences. However beyond the assertion that social welfare is a function of individual utilities it provides no guidance on the form of the function. As noted below it is questionable if even this assertion is correct. More seriously, unless the function is the unweighted summation of utilities it implies that the maximisation of utility will only coincidentally maximise social welfare.

However the objective in economic evaluation theory and practise is to maximise the unweighted summation of individual utilities (welfarism) or QALYs (extra-welfarism). This has been supported by the second approach to the problem of distribution. Welfare theory invokes the 'Kaldor-Hicks' or 'potential compensation principle'. One state is said to be superior to a second if there is the possibility of people in the first state compensating those in the second and still being better off. In the example above one QALY could be taken from each of the 3 people receiving 5 QALYs and given to the fourth person. The 3 QALYs would fully compensate the fourth person and leave the others better off. The compensation principle is, of course, inoperable: life years and QALYs cannot be transferred between people. A financial redistribution could be attempted to compensate losers. However, taxing winners – sick people who receive services from an NHS – conflicts with the purpose of an NHS. Compensating losers who have died because they received no service is not possible.

The attempt to dismiss or downgrade the importance of distributional issues is therefore problematical. As John Rawls notes

> '… distribution – indifference does not take the distinction between persons … seriously. If a person remains miserable or painfully ill, a deprivation is not obliterated or remedied or overpowered simply by making someone else happier or healthier. Each person deserves consideration as a person and this mitigates against a distribution – indifference view …' (p8) (quoted by Sen [22]).

Theory and methods which provide no guidance on how to approach the problem of distribution are incomplete and potentially inconsistent with social values.

Social preferences

The third problem arises from the assumption of individualism – that the correct perspective for assessing costs and benefits is only the perspective of the self-interested individual. However since the time of Aristotle it has been known that the preferences of the self-interested individual may differ from the preferences of the same individual in their role as a citizen [23]. As summarised in Box 1, decision making based upon an individualistic perspective and upon the social perspective of a citizen may differ. Self-interest may dominate in the former case but the citizen may be motivated by a wider range of considerations which are excluded from both welfarism and extra-welfarism by the assumption of individualism. To be sure, the perspective of the self-interested individual may be descriptively accurate for some people but it is an empirical question whether, on balance, people would prefer the alternative perspective to influence or dominate decision making in a particular context.

In the health sector a social perspective may result in a preference for the prioritisation of particular groups and for a distribution of benefits which differs from the distribution which would result from the maximisation of QALYs. These deviations may be subsumed under the heading of 'equity' and an equity-efficiency trade-off is recognised by Welfare Theory. However while the methods for achieving efficiency in the health sector and more generally in the economy, are developed and sophisticated, the methods for incorporating equity – community values and social preferences – are almost non-existent and most evaluation committees only consider clinical evidence and cost effectiveness [24]. At best, authorities such as the UK National Institute for Health and Care Excellence (NICE) which broadly support QALY maximisation modify the objective with adjustments to the cost/QALY threshold for some end of life treatments and for some highly specialised services for ultra-rare diseases. However, while the exceptions must fulfil a number of criteria, from the perspective of the economic theory supporting economic evaluation they are ad hoc. They confirm that the methods which result in the need for these changes and the theory upon which they are based are incomplete or inconsistent with social preferences.

Empirical evidence

Social values

Any theory or method which seeks to show how resources *ought* to be allocated must be based upon an ethical theory such as the utilitarian theory that utility *ought* to be maximised. This raises a further issue largely ignored by economic theory. Ethical rules cannot be derived from empirical evidence (the 'naturalistic fallacy' or Hume's 'is-ought' problem). Nevertheless decision makers are expected

Box 1. Social and individual perspectives in the health sector.
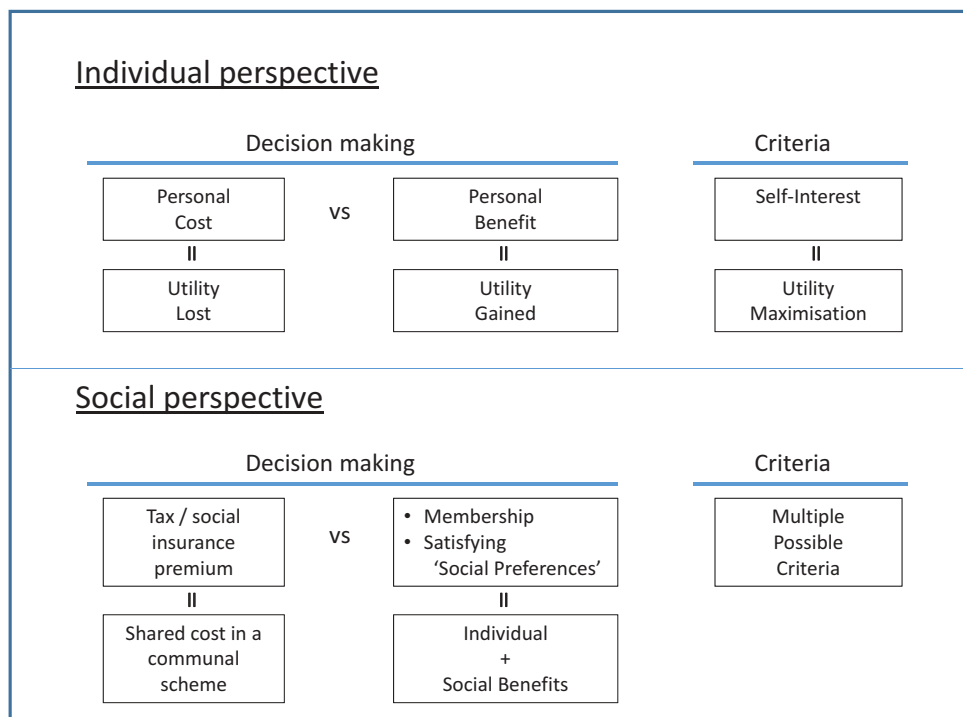
to act on behalf of citizens and subject to a number of caveats there is a strong case for communal values and social preferences – the preference of citizens – to be the basis for social policy, an approach which has been described as 'Empirical Ethics' [25–27]. Limited survey evidence exists which directly probe support for ethical principles. One such study was the 'Monash Health and Social Values Survey' of 455 representative Australians [28,29]. Results from five of its questions are reported in Table 1.

The response to questions 1 and 2 indicate that Australians are not hedonic utilitarians. They do not agree that the maximisation of happiness should be the sole ethical principle. The more significant finding is from the answers to questions 3–5. Less than 1 in 10 disagreed that they should fulfil their duties and that, setting aside the tautological interpretation of preference utilitarianism, this implies that the single criterion utility based weights used to calculate QALYs omits an important element or elements of social value.

Omitted attributes

A large number of empirical studies now exist which focus upon attributes omitted from the QALY model. Reviews of these have been published by a number of authors [20,30–32]. The most commonly identified attribute, illness severity, has been independently reviewed [30,33,34]. Other attributes include the patient's age and health potential, social status, type of service and illness. The uncertainty of future health also affects individual preferences but, as noted earlier, uncertainty has been largely ignored in the evaluation literature. A recent study by the first author found that 403 demographically representative Australians all selected more than the QALY maximising level of insurance for services for severe health states. Their selected insurance implied that the QALY maximising allocation of the budget under-valued these services by at least one third [16].

The significance of rare health states has recently become a controversial issue because of the increasing

Table 1. Results from the Monash social value survey (n = 455)[a.]

Statement	Agree	Disagree	Unsure
1. Action producing happiness is always right	22.8	57.4	19.8
2. Maximising happiness is more important than any other principle	14.3	65.9	27.8
3. I must fulfil duties even if it makes me less happy	92.1	7.9	0
4. having duties is a natural part of being a member of society	94.9	5.1	0
5. People help others only because they gain something personally	18.2	60.7	21.1

[a] Unpublished results

number of drugs and services for 'orphan disorders', services for ultra-rare diseases (URDs) [35–38]. Survey evidence indicates no preference for the special treatment of rare diseases when the choice is between rare and common diseases [39]. However as indicated by the results of the sharing surveys summarised below, these results are sensitive to the framing of questions. Rarity per se is not viewed as a reason for special treatment. But when the consequence of rarity is shown to be a very low cost per person, the framing of the question alters the response. This is also the conclusion of a recent, large scale survey of Swiss citizens when cost was presented as an increase in their insurance premiums [6,38,40–42].

Sharing surveys

Commencing with Nord et al [36] it was found that limited services would be shared with patients whose treatment was less cost effective than the treatment of an otherwise identical and clearly defined group of patients. The result has been replicated in several studies [35,43,44]. In each of these, however, it was assumed that treatment provided patients with a complete cure. With a fixed budget, this meant that a larger number of patients received no service.

In the four 'sharing surveys' described below, conducted by the first author, this constraint was removed and it was possible to provide a variable level of treatment. It was therefore possible to provide at least partial care to every patient. The cost effectiveness of different services was varied so that QALYs or life years could be maximised by allocating the fixed budget to services with the lowest cost per QALY. Other services would not, therefore, receive resources: sharing would not occur and some patients would be excluded from care. The surveys are summarised in Table 2.

Survey 1: Respondents were initially asked to allocate a budget to one of four patients who were identical except for the cost effectiveness of their treatment. The budget was increased up to 29 times and at each step respondents were asked to allocate the incremental budget to only one of the patients. At each step, the budget could increase the life of the four patients by 4, 6, 8 and 12 years respectively until they reached a limit of 48 additional years. The QALY maximising strategy was therefore to allocate each budget increment to the patient gaining 12 years until they reached their limit, then each increment to the patient gaining 8 years, then to the patient gaining 6 years and only then to the patient gaining 4 years. The average allocation of the 501 respondents to the survey is shown in Figure 1. In contrast with the QALY maximising strategy, after the

Table 2. Four sharing surveys.

Survey	Benefit	Study design: Allocation of a fixed budget which then increases	Key results	n	Reference
1	Life years (LY)	4 patients 3 fold variation in cost/LY	See Figure 2	501	Richardson et al [70]
2	Life years (LY)	Patient 2: cost/LY = 2 ×patient 1: condition less urgent	25% of total life years given to patient 2	430	Richardson et al [16]
3	QoL	Patient 2: cost/QALY = 3 ×patient 1: condition less severe	30% of total benefit given to patient 2	203	Richardson et al [71] under review
4	QoL	Group 2: cost/QALY = 20 ×group 1 Group 1: n = 100, 300, 600	Budget allocation to group 2: 26% of full cost when group 1 = 100 64% of full cost when group 2 = 600	432	Richardson et al [72]
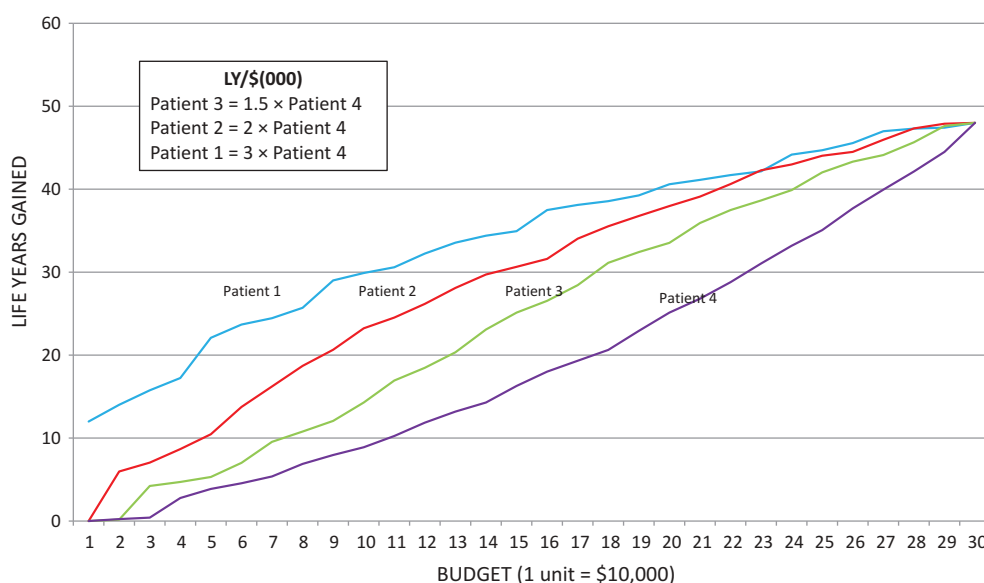
Figure 1. Results from sharing survey 1 (n = 501).

Source: Richardson et al. with permission from Elsevier [70]

first 2–3 stages respondents allocated the budget to patients who had received nothing.

In surveys 2–4 it was possible to divide the budget at each stage of the allocation exercise between two patients, patient 1 and patient 2. Only the key results are reported below.

Survey 2: In survey 2 the cost/life year for patient 2 rose to be double the cost for patient 1. Additionally, treatment was less urgent for patient 2: their life expectancy without treatment was 10 years: for patient 1 it was 2 years. Both QALY maximisation and prioritisation on the basis of urgency/need would therefore result in the full budget allocation to patient 1. This did not occur. After 6 increments to the budget patient 2 had been allocated an average of 25 percent of the life years and 20 percent of the potential QALY gain had been sacrificed to share the budget with patient 2.

Survey 3: The third survey largely replicated survey 2 except that the QoL replaced life expectancy as the main variable. Cost per increment of utility (QoL) for patient 2 rose to three times the cost for patient 1 while their QoL was 30 or 40 points higher on a 100 point visual analogue scale (VAS) in two parts of the exercise. QALY maximisation and/or prioritisation according to severity would therefore result in an allocation only to patient 1. In contrast, after 6 budget increments patient 2 had received an average of 30 percent of the total improvement in the QoL. In this study only 1 percent of respondents maximised QALYs and, on average, up to 65 percent of potential QALY gains were sacrificed in order to share the budget.

Survey 4: The QoL was also the main variable in the fourth survey. The difference in the cost effectiveness of services was significantly increased but the survey was

also designed to test the hypothesis that high cost services for rare diseases (SRDs) might be funded, not because they were rare per se but because, despite their high cost, rarity implies a low total cost and a small effect upon those bearing the cost. An analogy invoked was the sailor lost at sea whose rescue would cost $10 million which would be highly cost ineffective. But most would be prepared to pay the average cost per taxpayer of 50 cents rather than abandon the sailor.

The cost of incremental improvement in the QoL for group 2–5 patients – varied from 2 to 20 times the cost for group 1, which varied from 100 to 600 patients. QALY maximisation would result in no allocation to group 2. The 'sailor lost at sea' hypothesis suggests an increased allocation to group 2 as the number bearing the opportunity cost (a reduced QoL for patients in group 1) rose from 100 to 600.

The visual aid for the complex study design is shown in Figure 2. The vertical axis was calibrated using health states from the EQ-5D and placed to represent utilities found for these states in other surveys. The horizontal axis measured (to scale) the number of patients. The shaded areas therefore represented the QALY gain (persons × QoL) obtained by the two groups. As the slider at the bottom of the aid was moved left or right more

or less of the budget could be allocated to group 2 ('illness A'). An algorithm adjusted the shaded areas to meet the budget Results are shown in Figure 3. QALYs were not maximised in any scenario and resources were allocated to group 2 with all combinations of cost. As the size of group 1 increased and the effect of helping patients in group 2 was spread over more people, the allocation to group 1 increased. Up to 37 percent of potential QALYs were sacrificed to help patients in group 1.

Discussion

There are significant problems with the theory of economic evaluation and its policy prescription that QALYs should be maximised. In its measurable forms 'utility' does not take into account important individual preferences such as an aversion to uncertainty and the preference for greater protection against severe health states than provided when QALYs are maximised. Neither welfarism nor extra-welfarism distinguish the preferences of the selfish individual from the social preferences of the citizen and, through the use of the cost/QALY *ratio* as the decision criterion, the methods of economic evaluation rule out the potential
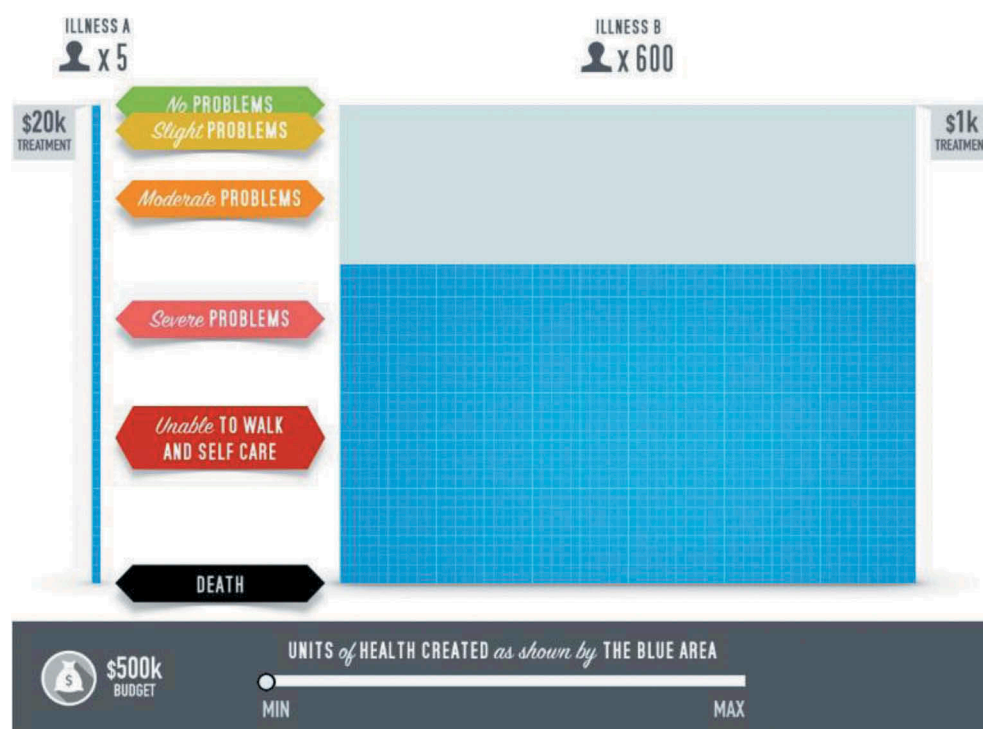

Figure 2. Visual aid for sharing study 4[(1)(2).]

(1) The shaded area indicates the 'health' obtained by the two groups. The vertical axis measures utility and the horizontal axis the number of patients affected. In this figure the 5 high cost patients have full health; the 600 low cost patients have a utility of 0.67. These could be changed by moving the slider at the base of the figure. The figure is to scale. The blue area therefore measures QALYs, although it was referred to as 'health' by the avatar.(2) Illness A was experienced by patients in Group 2; Illness B by those in Group 1.Source: Richardson et al [72]
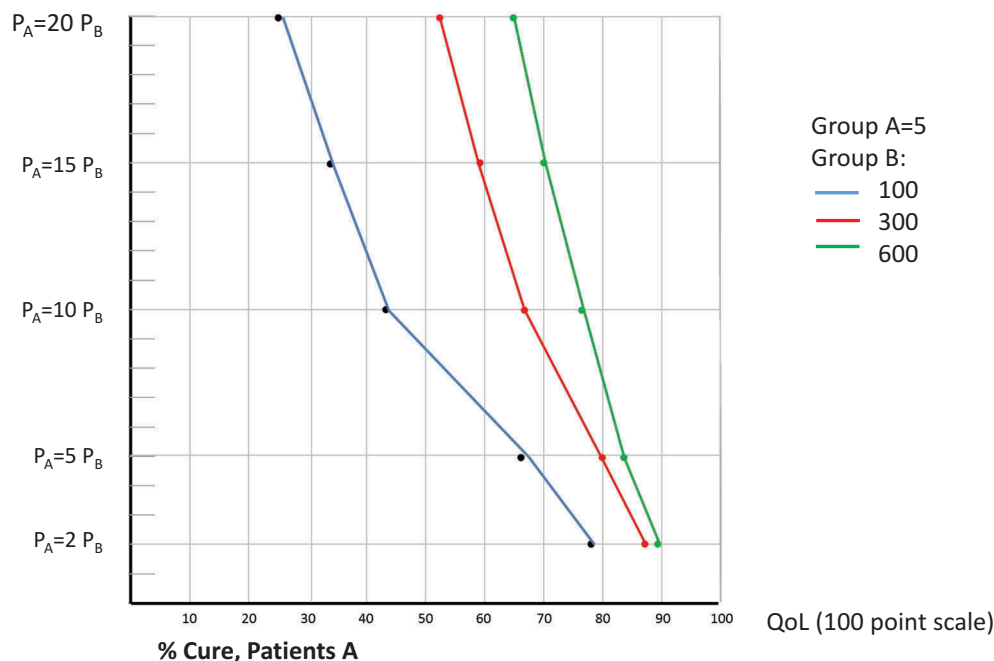
Figure 3. Results from sharing survey 4: QoL purchased for high cost patients A, [(1).]
(1) By price of insurance and size of the low cost group B, whose QoL falls to meet the budget.Source: Richardson et al [72]

importance of rarity. However social preferences are influenced by a variety of considerations excluded from the QALY model and, in particular, by the distribution of benefits. While QALY maximisation allows for the possibility that some individuals will receive no benefits evidence presented here suggests that the exclusion of patients from care conflicts with social preferences.

Exclusion from care also conflicts with the ethics of medical practise and generally does not occur irrespective of cost. It has been estimated for example that the cost of hospital, procedural and pharmaceutical services for accident victims in Australia in their final month of life in 2002 was $8,913 [45]. This translates into an annualised cost per QALY in 2018 of $US110,000. A more recent Australian study estimated that the monthly end of life cost of treating cancer patients in 2007 was $11,000 [46]. This translates to an annualised cost per QALY in 2018 of $US120,000. Despite wildly exceeding the de facto Australian threshold cost per QALY, it is doubtful if anyone would wish to reduce the care given to these patients.

It may appear improbable that the failures of theory and allocation methods described here would remain uncorrected. In part their persistence is a result of history. Huge analytical effort has been devoted to the development of these methods and their theoretical justification and for well-known reasons this inhibits change [47]. More importantly, the failures are largely immune to

empirical error learning. Flawed theory and methods are not directly observable: mistakes do not result in a bridge collapsing or a stock market crash. Rather, some people are treated inequitably and suffer silently.

One criticism of the empirical evidence presented here is that people's stated and revealed preferences may differ – that answers to a survey may be an unreliable guide to real world behaviour. However both behavioural economics and anthropology have demonstrated that sharing appears to be hard wired into people's behaviours [48–50]. In the evolutionary biology literature this is explained in terms of weak reciprocity or reciprocal altruism, that individuals will support the provision of benefits to others in the expectation that they would receive similar benefits if circumstances were reversed [51–53].

While this motivation may be consistent with self-interest it does not imply that the utilitarian framework as it is operationalised in the 'QALY model' is appropriate for economic evaluation. The selfish preferences of an individual motivated by reciprocal altruism would be similar to those of the citizen who is concerned with all outcomes as they are likely to view the world from behind a veil of ignorance – a recognition that they cannot foresee their future health or the services they will need. These include services which are not cost effective.

However people's behaviour is also consistent with strong reciprocity: people will reward or punish others at personal cost in the absence of the expectation of

Table 3. Elements of two competing paradigms.

Attribute	Extra-Welfarism	Communitarianism
Analytical Focus	Maximisation	Optimisation
Perspective on value	Consumer perspective (selfish 'use value')	Citizen's or social perspective (including risk aversion, caring, externalities and sharing)
Perspective on cost	Individual (patient)	Citizen (tax payer)
Social objective	Maximum 'health' defined by QALYs, ie utility weighted life years	Fair sharing: criteria based upon social preferences: reflecting equity and/or rights
Evaluation method	CUA CBA CEA	Social Cost Value Analysis
Criterion for funding	Cost/QALY < threshold,	Presumed entitlement
Funding formula	If criterion met, then services generally funding	Level of treatment varies with attributes such as rights/equity, cost effectiveness
Exclusions from care	Yes: Cost/QALY >Threshold	Few (except for extreme cases, usually milk and self-limiting health problems)
Caveat	Ad hoc adjustment for (to date) undefined 'equity'	Systematic adjustment, with budgetary impact and/or cost effectiveness per citizen
Role of cost	Pivotal: maximum benefit ← min cost/QALY	Secondary: alters the intensity of care
Ethical basis	Preference Utilitarianism	Communitarianism: satisfaction of social preferences

future gain when social norms are violated [54–57]). Evidence for this includes the ubiquitous results from the Ultimatum Game and its variants in which individuals forego a substantial financial gain to punish those who share benefits unfairly [10,11,58,59].

These established behaviours reinforce the evidence with respect to social values and sharing summarised here and suggest that prevailing preferences are better explained by communitarianism and rights based theories than by welfarism or extra-welfarism. These theories have been proposed as a superior basis for the allocation of health services, most notably by Mooney [60–62] but not developed into practical formula for allocating the budget.

The implications of this paradigm for economic evaluation are summarised in Table 3. The focus upon services and maximum efficiency defined by QALY gain is replaced by a focus upon patients and their fair treatment. This would include the presumption that all patients would be entitled to some level of care. To meet the health budget, the extension of care to include some cost ineffective procedures would require a reduction in the level of care for those who presently receive cost effective treatments. Results of the 'sharing' studies and Ultimatum Game indicate a public willingness to accept this redistribution to achieve a fair outcome. Finally, as emphasised in the communitarian literature, these outcomes are consistent with defensible ethical principles. Simple utilitarianism is neither the only ethical theory nor the most easily defended when it allows for an inequitable distribution of benefits and the possible exclusion of people from care.

The critique presented here implies the need for a redirection of research and a revision of funding criteria. The research agenda should have three foci; (i) additional research into social preferences and values; (ii) methodological research into the

instruments which measure social preferences; and (iii) practical research into the criteria or algorithms which will reform present funding formula or guidelines.

(i) The extension of CUA to include a wider set of social values has been described as 'social cost value analysis' [63]. It seeks to quantify the relative importance in different countries of the attributes which determine population preferences. But, in addition, it should include an enquiry into underlying population values which, without evidence, Welfare-Theory assumes to be utilitarian. Values guide the selection of attributes and the form in which they should be quantified. Without the recognition of communitarian values sharing would not have been investigated. When population support for the funding of high cost orphan products is tested, a communitarian focus suggests the need to highlight the sharing of cost across the insured population.

(ii) There is a two-fold challenge in the measurement of social values. First multiple attributes must be considered and combined. Second, the risk of measurement invalidity due to cognitive overload must be overcome. To date measurement has primarily used the Person Trade-Off (PTO) instrument (cf review articles cited earlier). The Relative Social Willingness to Pay (RS-WTP) was introduced by the first author to overcome a number of its perceived weaknesses [64] and more recently Discrete Choice Experiments (DCE) have been employed. To date the former two techniques have been used to measure the importance of single attributes. In the case of the first author's surveys they have focussed upon the need to minimise the cognitive burden of complex decisions. DCE seeks trade-offs between multiple attributes. It simplifies decision making by requiring binary choices. However when the number of attributes is large it is questionable whether respondents have the cognitive capacity to simultaneously appreciate all of the attributes and their

implications and to refrain from simplifying heuristics such as a disproportionate focus upon a single attribute. DCEs are prone to a large number of potential biases, and it is difficult in complex experiments to demonstrate that each attribute is considered equally [65].

Options for future measurement have been reviewed elsewhere [63]. They include Multi-Criteria Decision Analysis (MCDA), a technique which allows the combination of independently evaluated attributes [66,67]. The existence of multiple measurement techniques indicates the need for research into their reliability and validity. A salutary lesson with respect to the need for this research could (but has not) been learned from the measurement of utility. Instruments purporting to measure the same quantity – utility – give significantly different results with respect to the importance of different health problems [68].

(iii) Research, into the reform of funding formula, has scarcely commenced. Nord has suggested how severity could be incorporated into decision making [69]. In the fourth sharing study reviewed earlier there is a funding formula which illustrates how a fixed budget may be allocated between services with a 20-fold difference in the cost/QALY while taking account of cost and severity. However formula incorporating multiple attributes and empirical parameters remain a work in progress. Consequently, and contingent upon further confirmatory studies, the social preferences identified to date can only be accommodated with discretionary judgements. These should, minimally, be explicit and include a recognition of the independent importance of severity [30,69] and the need to fund 'cost ineffective' services when alternative treatments are not available.

Conclusions

Technological advances have resulted in an increasing number of high cost services indicated for rare, and often severe disorders. Applying conventional health economic evaluation criteria would lead to the exclusion of most of these services from funding, because they do not meet commonly used benchmarks for cost effectiveness. However, the evidence reviewed here and elsewhere indicates that this outcome would be inconsistent with prevailing social values. The inconsistency is attributable to shortcomings in the theoretical foundations of the evaluation paradigm. There is, therefore, a need for further research into social preferences and the means for translating them into workable funding formulae.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Jeff Richardson http://orcid.org/0000-0002-4248-4280
Michael Schlander http://orcid.org/0000-0002-3489-5856

References

[1] Culyer A. The normative economics of health care finance and provision. Oxford Rev Econ Policy. 1989;5 (1):34–58.

[2] Brazier J, Ratcliffe J, Salomon J, et al. Measuring and valuing health benefits for economic evaluation. Second ed. Oxford: Oxford University Press; 2017.

[3] Rice T. The economics of health reconsidered. Chicago, IL: Health Administration Press; 1998.

[4] Bentham J. An introduction to the principles of morals and legislation. In: Mill JS, Bentham J, editors. Utilitarianism and other essays. Harmandsworth: Penguin; 1824/1987.

[5] Edgeworth FY. The hedonic calculus. Mind. 1879;4:394–408.

[6] Schlander M. Measures of efficiency in healthcare: QALMS about QALYs? Z Evid Fortbild Qual Gesundhwes. 2010;104(3):214–226.

[7] Richardson J. Cost utility analysis: what should be measured. Soc sci med. 1994;39(1):7–21.

[8] Ng Y-K. From preference to happiness: towards a more complete welfare economics. Soc Choice Welfare. 2003;20:307–350.

[9] Kahneman D, Wakker P, Sarin R. Back to Bentham? Explorations of experienced utility. Q J Econ. 1997;112:375–405.

[10] Oosterbeek H, Sloof R. van de Kuilen G. Cultural differences in Ultimatum Game experiments: evidence from a meta-analysis. Exp Econ. 2004;7(2):171–188.

[11] Sanfey AG, Rillinkg JK, Aronson JA, et al. The neural basis of economic decision-making in the Ultimatum Game. Science. 2003;300(5626):1755–1758.

[12] Anderson E. Beyond *homo economicus*: new developments in theories of social norms. Philosophy Public Affairs. 2000;29(2):170–200.

[13] Camerer CF, Thaler RH. Anomalies: ultimatums, dictators and manners. J Econ Perspect. 1995;9(2):209–219.

[14] Engel C. Dictator games: a meta study. Exp Econ. 2011;14 (4):583–610.

[15] Guala F, Mittone L. Paradigmatic experiments: the dictator game. J Socio Econ. 2010;39(5):578–584.

[16] Richardson J, Iezzi A, Maxwell A. Uncertainty and the undervaluation of services for severe health states in cost utility analyses. Value Health. 2017;21(7):850–857.

[17] Schoemaker P. The expected utility model: its variants, purposes, evidence and limitations. J Econ Lit. 1982; XX:529–563.

[18] Culyer A. Commodities, characteristics of commodities, characteristics of people, utilities and quality of life. In: Baldwin S, Godfrey C, Propper C, editors. Quality of life:

perspective and policies. London: Routledge; 1990. p. 9–27.

[19] Richardson J, Maxwell A, Khan MA, et al. The myth of maximisation: quantifying trade-offs between social objectives using the conflict scale, Research Paper 74. Melbourne: Centre for Health Economics, Monash University; 2012.

[20] Dolan P, Shaw R, Tsuchiya A, et al. QALY maximisation and people's preferences: A methodological review of the literature. Health Econ. 2005;14(2):197–208.

[21] Reinhardt U. Abstracting from Distributional Effects: this policy is efficient. In: Barer M, Getzen T, Stoddart G, editors. Health, health care and health economics (pp. 1-52). Chichester: John Wiley & Sons; 1998.

[22] Sen A. Why Health Equity. Keynote Address. Third International Conference on The Economics of Health Within and Beyond Health Care; 2001 23 July; New York. p. 13.

[23] Crisp R. Aristotle Nicomachean Ethics. Cambridge: Cambridge University Press; 2000.

[24] Ghijben P, Gu Y, Lancsar E, et al. Revealed and stated preferences of decision makers for priority setting in health technology assessment: a systematic review. Pharmacoeconomics. 2018;36(3):323–340.

[25] Richardson J, McKie J. Empiricism, ethics and orthodox economic theory: what is the appropriate basis for decision making in the health sector. Soc sci med. 2005;60:265–275.

[26] Hausman D. The limits to empirical ethics. In: Murray CJL, Salomon JA, Mathers CD, et al., editors. Summary measures of population health. Geneva: World Health Organization; 2002. p. 641–646.

[27] Richardson J. The poverty of ethical analyses in economics and the unwarranted disregard of evidence. In: Murray CJL, Salomon JA, Mathers CD, et al., editors. Summary measures of population health. Geneva: World Health Organization; 2002. p. 628–640.

[28] Anderson M, Richardson J, McKie J, et al. The relevance of personal characteristics in health care rationing: what the Australian public thinks and why. Am J Econ Sociology. 2011;70(1):131–151.

[29] Anderson M, Richardson J, McKie J, et al. Queues, private health insurance and medicare: is economics the driving force? Research Paper 23. Melbourne Centre for Health Economics, Monash University; 2007.

[30] Gu Y, Lancsar E, Ghijben P, et al. Attributes and weights in health care priority setting: a systematic review of what counts and to what extent. Soc sci med. 2015;146:41–52.

[31] Schwappach DL. Resource allocation, social values and the QALY: a review of the debate and empirical evidence. Health Expectations. 2002;5(3):210–222.

[32] Stafinski T, Menon D, Marshall D, et al. Societal values in the allocation of healthcare resources: is it all about the health gain? Patient. 2011;4(4):207–225.

[33] Nord E, Johansen R. Concerns for severity in priority setting in health care: a review of trade-off data in preference studies and implications for societal willingness to pay for a QALY. Health Policy. 2014;116(2):281–288.

[34] Shah KK. Severity of illness and priority setting in healthcare: a review of the literature. Health Policy. 2009;93:77–84.

[35] Abellan-Perpinan J, Pinto-Prades J. Health state after treatment: a reason for discrimination? Health Econ. 1999;8:701–707.

[36] Nord E, Richardson J, Kuhse H, et al. Who cares about cost? Does economic analysis impose or reflect social values? Health Policy. 1995;34(2):79–94.

[37] Schwappach DL. Does it matter who you are or what you gain? An experimental study of preferences for resource allocation. Health Econ. 2003;12:255–267.

[38] Schlander M, Garattini S, Holm S, et al. Incremental cost per quality-adjusted life year gained? The need for alternative methods to evaluate medical interventions for ultra-rare disorders. J Comp Eff Res. 2014;3(4):399–422.

[39] Ryan M, Scott DA, Reeves C, et al. Eliciting public preferences for healthcare: A systematic review of techniques. Health Technol Assess. 2001;5:1–186.

[40] Schlander M, Garattini S, Kolominsky-Rabas P, et al. Determining the value of medical technologies to treat ultra-rare disorders: A concensus statement. Market Access Health Policy. 2016;4:pp 1–9.

[41] Schlander M, Telser H, Fischer B, et al. Should rarity matter? Results of a social willingness-to-pay study using Discrete Choice Experiment (DCE) method in Switzerland. Value Health. 2018;21(S2):(in press).

[42] Schlander M Health economic evaluation of medical interventions: answering questions people are unwilling to ask. 5th World congress International Health Economics Association (iHEA), Book of Abstracts. 2005 July 10–13; Spain: Barcelona: 344–355.

[43] Ratcliffe J. Public preferences for the allocation of donor liver grafts for transplantation. Health Econ. 2000;9:137–148.

[44] Ubel P, Loewenstein G. Distributing scarce livers: the moral reasoning of the general public. Report. Pennsylvania: Veterans Affairs Medical Centre; Division of General Internal Medicine, University of Pennsylvania; Centre for Bioethics, University of Pennsylvania; 1995.

[45] Kardamanidis K, Lim K, Da Cunha C, et al. Hospital costs of older people in New South Wales in the last year of life. Med J Aust. 2007;187(7):283–386.

[46] Reeve R, Srasuebkul P, Langton JM, Haas M, Viney R, Pearson S-A, et al. Health caer use and costs at the end of life: A comparison of elderly Australian decedents with and without a cancer history. BMC Palliat Care. 2017;16(1):1–10.

[47] Kuhn TS. The structure of scientific revolutions. Chicago: Chicago University Press; 1962.

[48] Bowles S, Gintis H. A cooperative species: human reciprocity and its evolution. J Econ Issues. 2011;46(2):575–577.

[49] Gintis H. Altruists together: herbert Gintis applauds two books that powerfully enrich the dialogue on behavioural science. Nature. 2015;517(7536):550.

[50] Kameda T, Takezawa M, Tindale RS, et al. Social sharing and risk reduction: exploring a computational algorithm for the psychology of windfall gains. Evol Human Behav. 2002;23(1):11–33.

[51] Falk A, Fischbacher U. The economics of reciprocity: evidence and theory. In: Freeman R, editor. Inequality arond the World. UK: Palgrave Macmillan; 2002. p. 207–233.

[52] Fehr E, Schmidt KM. The economics of fairness, reciprocity and altruism - experimental evidnece and new theories. InKolm SC, Ythier JM editors. Handbook of the economics of giving, altruism and reciprocity. North

Holland: Elsevier; 2006. p. 615–691. DOI:10.1016/S574-0714(06)01008-6

[53] Pelaprat E, Brown B. Reciprocity: understanding online social relations. First Monday. 2012;17(10). DOI:10.5210/fm.v17i10.3324

[54] Fehr E, Fischbacher U, Gächter S. Strong reciprocity, human cooperation and the enforcement of social norms. Hum Nature. 2002;13:1–25.

[55] Fraser B. Explaining strong reciprocity: cooperation, competition, and partner choice. Biol Theory. 2011;6 (2):113–119.

[56] Guala F. Reciprocity: weak or strong? What punishment experiments do (and do not) demonstrate. Behav Brain Sci. 2012;35:1–59.

[57] Trivers RL. The evolution of reciprocal altruism. Q Rev Biol. 1971;46:35–57.

[58] Andersen S, Ertac S, Gneezy U, et al. Stakes matter in Ultimatum Games. Am Econ Assoc. 2011;101(7):3427–3439.

[59] Nicklisch A, Wolff I. On the nature of reciprocity: evidence from the ultimatum reciprocity measure. J Econ Behav Organiz. 2012;84(3):892–905.

[60] Mozoney G. Communitarian claims as an ethical basis for allocating health care resources. Soc sci med. 1998;47 (9):1171–1180.

[61] Mooney G. Communitarian claims and community capabilities: furthering priority setting? Soc sci med. 2005;60 (2):247–255.

[62] Mooney G. Challenging health economics. Oxford: Oxford University Press; 2009.

[63] Caro JJ, Brazier J, Karnon J, et al. Detemining value in Health Technology Assessment: stay the course or tack away? Pharmacoeconomics. 2018 Nov 9.

[64] Richardson J, Iezzi A, Sinha K, et al. An instrument for measuring the relative social willingness to pay for health state improvement. Health Econ. 2013;23:792–805.

[65] Kjaer T A review of the discrete choice experiment - with emphasis on its application in health care, (Health Economics Papers; No. 1). Denmark: University of Southern Denmark; 2005.

[66] Marsh K, Laniti T, Neasham D, et al. Assessing the value of healthcare interventions using multi-criteria decision analysis: a review of the literature. Pharmacoeconomics. 2014;32(4):345–365.

[67] Wahlster P, Goetghebeur M, Kriza C, et al. National leading-edge cluster medical technologies 'medical valley EMN'. Balancing costs and benefits at different stages of medical inovation: systematic review of multi-criteria decision analysis (MCDA). BMC Health Serv Res. 2015;15(262):pp1–12.

[68] Khan MA, Richardson J. Variation in the apparent importance of health related problems used to measure patient welfare. Qual Life Res. 2018 August 18.

[69] Nord E. Cost-Value Analysis of health interventions: introduction and update on methods and preference data. Pharmacoeconomics. 2015;33:89–95.

[70] Richardson J, Sinha K, Iezzi A, et al. Maximising health versus sharing: measuring preferences for the allocation of the health budget. Soc Sci Med. 2012;75(8):1351–1361.

[71] Richardson J, Iezzi AMaxwell A. Sharing and the provision of "cost ineffective" life extending services to less severely ill patients. Value In Health. 2018;21:951-957. DOI: 10.1016/j.jval.2017.12.024

[72] Richardson J, Iezzi A, Chen G, et al. Communal sharing and the provision of low volume high cost health services: results of a survey. Pharmacoeconomics. 2017;1:13–23.