

Methods of Economic Evaluation: What are the Ethical Implications for Healthy Public Policy?

November 2014

Introduction: The ethical implications of economic evaluations

Decision making in healthy public policy,¹ as in all policy areas, increasingly involves taking economic efficiency into consideration. Efficiency is the extent to which sought-after benefits can be obtained for the lowest possible cost, and the tools that measure it are economic evaluations. Efficiency is, however, but one of the many possible criteria according to which policy options can be judged. There is a range of other values and objectives that we may want policies to fulfill. Deciding between at times divergent values is an ethical enterprise, and the use of economic evaluations can have profound ethical implications.

The first paper in this series introduced some of the general ethical issues that arise when economic evaluations are applied in healthy public policy.² While there are a number of diverse methods of economic evaluation, all of them share several fundamental, underlying assumptions that have ethical implications. Most prominent among these are the assumptions of individualism in methodology and utilitarianism in ethics. Methodological individualism is the assumption that, simply put, all “we’s” can be reduced to collections of “I’s”; in other words, all social phenomena can be explained with reference solely to the actions and beliefs of individual human beings. Such an assumption tends to promote values such as individual autonomy and can conflict with values such as social solidarity and community empowerment that are based on a more holistic understanding of communities.

The second major assumption, utilitarianism, defines good and bad purely in terms of “utility.” In the version of utilitarianism most common in economics, more utility stems from that which people would prefer given a choice between several options. The more people’s preferences are satisfied, the better; hence, this version is often called the “preference-satisfaction view.” As with individualism, there are important ethical ramifications stemming from the assumption of utilitarianism. For example, almost any value judgment can be termed good – an individual only has to prefer it. However, if we imagine, as we sometimes do in public health, that some preferences can be conditioned and may actually be harmful, then a conflict can arise. Another series of issues comes from the maximizing nature of utilitarianism: its main goal is to maximize the number of satisfied preferences, not to reach a certain distribution of satisfied preferences. In a conflict between the two broad goals of public health, the maximization of health gains for the whole population and the reduction of health inequalities between subpopulations, economic evaluations will prioritize the first goal over the second, thus potentially leading to recommendations that will tend to exacerbate inequalities.

Although economic evaluations are evidence-informed and use hard numbers to gauge efficiency and, as such, their results may seem to be similarly hard facts, they are not value-free. At the most fundamental level, they are based on assumptions that can lead to significant ethical questions in policy decision making.

Yet such ethical questions do not only stem from these deep-down assumptions. Economic evaluations vary according to the question under study and the methods they use; each method

¹ Healthy public policies are policies that usually fall outside of the scope of the health sector, but which can nonetheless have important benefits for the health of the population while pursuing other aims. Examples of healthy public policies can include social housing policies, traffic-

calming policies, zoning bylaws to restrict the number of fast-food outlets near schools, etc.

² This first paper is available here: http://www.ncchpp.ca/144/Publications.ccnpps?id_article=962



also has its own, specific ethical implications. The aim of this paper is to highlight the ethical issues that arise from the differences between methods of economic evaluation. Although these methods broadly share the key assumptions just described, they differ substantially in their specifics, most notably in how they calculate effects.

The various methods of economic evaluation for public policy generally, and healthy public policy specifically, all seek to find the most economically efficient policy option to pursue. This usually means determining the option with the most “bang-for-the-buck.” To locate this best choice, it is necessary to determine the costs of competing policy options as well as their effects – in other words, to gauge how the relationship between costs and effects is distributed across alternatives, including, as always, the option of maintaining the status quo. This paper begins by looking at how costs are determined across the various methods of economic evaluation. Costing, while not trivial, has ethical issues that are largely shared by all methods. The differences between methods, on the other hand, truly emerge when looking at how they measure and value the effects stemming from policy options. Indeed, the question of defining and measuring effects is often the harder one and the one that has broader ethical implications. It will be discussed at length in the third section with a breakdown by method. The fourth section will focus on the important issue of whose point of view is the appropriate one to adopt in decision making. The paper will conclude with some thoughts on method choice and what it might mean to use the right tool for the job in a policy context.

List of the main acronyms

CBA	Cost-benefit analysis
CCA	Cost-consequence analysis
CMA	Cost-minimization analysis
CUA	Cost-utility analysis
MCDA	Multi-criteria decision analysis
QALY	Quality-adjusted life year
WTP	Willingness-to-pay

Calculating costs

Regardless of the method used to evaluate a policy, costs and effects are somehow compared. While there are numerous approaches to treating effects, costs are universally analyzed in dollar terms. This does not, however, make costing an easy or ethically neutral process. Indeed, the resources to be used in each policy under consideration have to be correctly identified, measured and valued (Brouwer, Culyer, van Exel, & Rutten, 2008) and each of these steps provides its own set of concerns.

A. IDENTIFYING RESOURCES – OR WHAT KINDS OF RESOURCES?

The first challenge in costing a policy option often lies in simply correctly identifying costs. In general, costs are any resources necessary to bring a policy option about. They include outlays as well as ongoing expenses; they can be explicit or sometimes merely implicit. Further, some evaluations can include all negative consequences as costs.

The breadth of healthy public policy means that the resources interventions utilize are not always limited to one ministry, one level of government or one agency nor do they often arise solely from direct health care (Simoens, 2009). A full list of resources to be costed can include not only the physical materials and staff time required but also items such as productivity gains or losses, time diverted by family and caregivers and much more. For example, a recreation program for seniors might require equipment purchases, updates to existing infrastructure, new staff positions for activity leaders as well as transportation and caregiver time costs associated with bringing participants to and from the program.

It may be the case, whether explicitly or implicitly, that only some costs are taken into account in an evaluation. Such selectivity may, for example, stem from narrowing the perspective to include only those costs borne by a particular agency; the issue of perspective will be discussed in more detail on pages 13-14. Taking into account only some costs while ignoring others is not necessarily problematic; it may be desirable, depending on the situation. If we want to positively discriminate in favour of a certain population, for example, it may be relevant to ignore some of the costs borne by the rest of society and focus only on those borne by this group in the analysis. At other times, however, a more

circumscribed point of view can be detrimental – an incomplete accounting of costs can lead to a less than optimal policy choice. For example, if not all costs are taken into account, then policies can appear to be more beneficial than they would have been, had a comprehensive cost analysis taken place.

B. MEASURING RESOURCES – OR HOW MANY RESOURCES OF EACH KIND?

Once resources have been correctly identified, they have to be measured. Here, the potential issues are more methodological than ethical. Different methods of measurement can produce different cost estimates. Measures can use different scales; for example, there can be differences between the macro-level total cost of a service and more micro-level costs like that of service cost per user, which can vary depending on anticipated users. Measures can also vary depending on their source. Whether a cost analysis is based on past cost studies, current statistics or best prospective estimates matters and can lead to different outcomes (Brouwer et al., 2008). Finally, extrapolating costs further and further into the future also poses difficulties as uncertainty – both about the size of costs and whether particular costs will be actualized – grows the further we project into the future. A full accounting of the ethical implications of uncertainty is too broad a topic to include in this paper.

C. VALUING RESOURCES – OR HOW MUCH ARE THEY WORTH?

Having identified and measured resources, the final task in preparing a cost analysis for an economic evaluation is to place a value on these resources. In economic evaluations it is typical to value resources at their “opportunity cost” as per standard economic theory (Simoens, 2009). Opportunity cost refers to the value a resource has in its next best alternative use. Standard economic theory assumes that in a free and competitive market, opportunity costs are reflected in market prices. Indeed, many items, from signposts to entire buildings, can be bought on the market and are thus relatively easy to cost. Other things, however, are much harder to accurately value and for this reason may not always be included. As discussed earlier, depending on the context, this kind of omission can have important ethical implications.

Time and care are two examples of resources that do not have well-defined market prices. For example, time and productivity lost at work can be included in some health-related economic evaluations. While an individual’s wage is seen to be a good proxy for this, it is not always clear what wage to use. Options are plentiful and so, difficult questions abound. Should one use the average national wage or a regional wage, the mean or the median? Perhaps instead one should use the average wage of the particular demographic at which a program is aimed? If so, how is this group defined and how does this skew the cost of a program? These are some of the questions that may arise. There is also the question of those who do important work but do not get paid, such as informal caregivers, who are predominantly women (Brouwer et al., 2008). It is important to consider how their time should be included and how it should be valued – especially since this increases program costs and so may be only reluctantly included.

These are not merely technical problems; they reflect deep ethical questions about how we value people, their energy and time. On the one hand, we may want to value all people and their time equally simply on the basis of the principle of equality. On the other hand, we may be more interested in equity and targeting disadvantaged groups. Here, using the actual valuations of lower-income individuals’ time will lower costs and potentially make policies directed toward them more attractive in comparison with other options targeted at the general population or other groups. Similarly, if we want to ensure that policies take into account issues of gendered work (both in and outside the household), then valuing caregiver time appropriately can be a good starting place and open up space for taking concrete steps to value caregivers during policy implementation.

Calculating costs – Summary

To calculate costs, one must:

- a) Identify all types of relevant resources
- b) Measure the quantity of each type of resource needed
- c) Evaluate how much these resources are worth

Questions to keep in mind to identify ethical implications:

- Are all relevant costs included (kinds of resources and quantity of each kind)?
- Are the ways the resources are valued, especially those without well-defined market prices (people's time, for instance), at risk of skewing the results in favour of interventions benefiting and/or burdening some subpopulations?

Both the relevant direct benefits – those factors which the policy in question aims explicitly to improve – and indirect effects – those factors on which the policy in question has a positive or negative effect but which are not among its stated aims, such as productivity gains – should be equal. Only in such a situation can benefits be safely ignored and the analysis focus merely on costs. An example where CMA might be appropriate is a decision over whether to use public or private procurement for the delivery of a given project, whether it be several kilometres of bike path or a new health facility. In such situations, what counts as the most efficient course of action is the lowest-cost means that achieves the desired end.

Nevertheless, there may be ethical implications stemming from the use of CMA even in such seemingly clear-cut situations. In the decision between public and private procurement, assuming that benefits are equal because the end product is the same can omit important potential indirect effects. For example, public procurement can create and maintain expertise within public administrations that can later aid in other public projects. The question of how much value we place on such expertise is itself an ethical one, but one that cannot be posed within the bounds of a CMA. More complex varieties of evaluation make space for such indirect benefits to be potentially uncovered and the resulting broader ethical questions to be considered.

Calculating benefits and other effects: The different approaches to economic evaluation

The other side to determining the costs of a policy is, of course, calculating the benefits it will produce if implemented. It is here that the various methods of economic evaluation truly differentiate themselves. Like costs, benefits have to be identified, measured and valued; unlike costs, however, there are myriad ways to accomplish these tasks, particularly the task of assigning value to benefits. The four main methods of economic evaluation will be described below with a focus on the ethical implications arising from each.

A. **COST-MINIMIZATION ANALYSIS (CMA), OR LEAVING BENEFITS ASIDE**

One of the oldest methods of economic evaluation is cost-minimization analysis (CMA). It is also the simplest method because it does not require benefits to be calculated. In order for this to be possible, CMA can only be applied when benefits are assumed to be nearly uniform, both in scope and in nature, across the policy options being considered. Indeed, the most important step in performing a CMA is to determine whether it is appropriate for the given situation: it is crucial that the expected outcome of whatever policy options are being considered be essentially the same (Jefferson, Demecheli, & Mufrod, 2000).

Cost-minimization analysis (CMA) – Summary

Ranks options deemed to have the same direct and indirect effects according to their costs. The least costly option is deemed the more efficient.

A question to keep in mind to identify ethical implications:

- Do the options compared really have the same direct and, especially, indirect positive and negative effects?

B. **COST-BENEFIT ANALYSIS (CBA), OR MONETIZING BENEFITS**

Few policy decisions are so clear-cut as to warrant CMA, and more nuanced methods of economic evaluation have been devised and put to use, including in decision making for healthy public policy. An important method, developed early in the history of modern economics but still maintaining notable

popularity, is cost-benefit analysis (CBA). Indeed, although CBA is what many people associate with all economic evaluations, it is but one, concrete method with its own defining features, strengths, limitations and ethical implications.

Defining features

CBA is firmly grounded in traditional welfare economic theory – this means that, in short, it seeks the maximization of net gains in social well-being as measured by satisfied individual preferences. Its key assumption is that not only all costs, but all effects of an intervention can be expressed in monetary terms. Once costs and benefits are translated into dollar amounts, policy recommendations boil down to seeing which policy option under consideration results in the highest ratio of benefits to costs.

Expressing benefits in dollar terms, however, is not an easy task. The primary difficulty lies in finding a means to translate the multitude of possible, otherwise incommensurable benefits and other effects that a policy can generate – things as disparate as improvements in health, greater employment, better public safety and so on – into dollar amounts. To deal with this issue, evaluations have to choose how many effects to “monetize.” One question is which effects to include: on one end of the spectrum is a choice to directly monetize the less controversial effects and leave more difficult-to-value effects out of the equation. On the other end lies translating all effects into monetary terms. Many evaluations fall somewhere in between.

A second, more challenging question is how to come up with those monetary valuations. Again, there are two broad options. One method is commonly called “revealed-preference valuation.” It uses existing data on prices and behaviours to glean information about how people value things. For example, gauging the value of peace and quiet due to living in a traffic-calmed neighbourhood might involve calculating the average price difference of similar homes in two neighbourhoods – one traffic-calmed and one not – that are otherwise very similar.

The other option is referred to as “contingent valuation” and frequently makes use of the willingness-to-pay (WTP) framework. With WTP, individuals are asked how much they would be willing to pay for particular benefits and the answers to such questions are used to generate valuations. Unlike in the revealed-preference method, here the

dollar amount that a benefit is worth can be elicited directly. There are many ways of doing this, from asking explicitly how much one would pay for the benefit to engaging individuals in bidding games, asking them to choose between monetary gambles involving the benefit in question or using other proxies from which dollar values can be derived. For example, imagine wanting to determine the subjective value of being able to walk to work. Such a valuation could, for instance, be useful in evaluating the benefits of urban planning policies. One way of eliciting WTP would be to ask how much work time someone would be willing to forgo to be able to walk to work (described as, say, the time it took the employer to relocate to the new, walkable location). Here, the proxy is the individual’s wage rate and multiplying it by the time potentially forgone gives a dollar value. Another means of obtaining the same information would be via a bidding game: the individual is asked whether she would hypothetically pay \$X for being able to walk to work; if she answers in the affirmative, the same question is asked with a higher bid, if she answers in the negative, the bid is lowered. This continues for several rounds to reach a more precise figure.

Strengths

CBA has important strengths, especially as it can be applied to the evaluation of healthy public policy. Chief among these is universality: CBA could potentially be used in contexts where the choice is between policies that come from completely different fields, for example those that are clearly health-oriented and others that lack health-promoting components (Jefferson et al., 2000) – albeit this is not frequently done in practice. As such, CBA evaluations can be used to advocate for healthy public policies on efficiency grounds in policy areas where many other options may not be targeted, even partially, at health improvements. Cost-benefit analyses give healthy public policy advocates the potential to find a common language with their peers across a wide range of fields.

Another strong suit of CBA is its flexibility to handle any and all types of benefit (Schlander, 2010). Any benefit potentially resulting from a policy that does not have a well-defined market value can regardless be expressed in monetary terms using the various tools that elicit willingness-to-pay. Thus, benefits as disparate as improvements in air quality and increased protein intake can be compared on the same metric. As the benefits stemming from many

healthy public policies can be quite diverse, this flexibility is all the more an asset for this particular class of policies. Even critics of the WTP framework on which CBA is based admit that, particularly in today's conditions of scarcity of public resources, CBA is a tool that can help policy makers wisely earmark every valuable dollar across a range of competing and vastly different uses (Cookson, 2003).

Limitations

The major limitation of CBA stems, in many ways, from the same facts of its universality and flexibility that are its greatest strengths. It is hard enough to enumerate all the tangible and intangible benefits a policy may bring about; translating all of these into dollar terms adds an entirely new layer of difficulty. This is especially so for benefits that do not have well-defined market prices. Gauging their value can potentially introduce inaccuracies and biases. It is not always clear how to design the tools to obtain valuations, whom to ask and how to interpret the results, in particular for WTP-based methods.

Ethical implications

Indeed, the WTP framework raises a number of ethical issues (Cookson, 2003) when applied to healthy public policy. The first issue often raised is that WTP is closely tied to income levels (Gold, Siegel, Russell, & Weinstein, 1996). Critics believe that rather than measuring one's **willingness** to pay for health-related improvements, benefits as they are valued under CBA instead are more a measure of one's **ability** to pay for them. The same health outcome or health determinant may be given a much lower dollar value by a low-income individual than it would be by a wealthy individual merely on the basis of a bias stemming from different income levels. In general, monetary valuations generated by WTP may reflect income and wealth levels in addition to the importance attributed to them by those questioned. In practice, this means that using CBA based on WTP to compare policy options will tend to rank as more efficient options that primarily benefit the wealthy compared to options that primarily benefit less well-off individuals.

Aside from income, other biases can also potentially impact WTP valuations (Cookson, 2003). "Scope" effects can inflate the value of relatively small benefits. Studies have shown that people often appear to value reductions in harm at a similar, standard amount, regardless of the size of the

reduction (Cookson, 2003). Thus, harm reductions that lead to the extension of life by, respectively, a year and a decade could receive very similar dollar values in two different studies, even from the same respondents. "Budget constraint bias," on the other hand, leads people to value the same benefit more highly when it is presented in isolation than if it were presented as one among several other competing options. These and other similar biases together constitute "framing" effects – in short, the context, or frame, within which something is presented matters to our understanding and evaluation of it.

Framing is an ethical challenge for all valuation methods. Under revealed preference valuation, it simply manifests itself differently. For example, there is a fundamental question about the extent to which differences in so-called **shadow prices** (those real-world prices used as proxies for things without market prices) actually conform to differences in values (Wegner & Pascual, 2011). That is, the implicit framing of a market transaction may not allow for an expression of the full range of possible motivations behind valuation and choice. A related ethical consideration is that revealed-preference valuation takes as given the current distribution of goods. As such, the benefits of policies that also impact strongly on distribution may not be accurately reflected in revealed-preference studies (Hausman, 2011). For example, a housing program that impacts on real estate prices in a broader region will also change the shadow prices that could have been used to justify it.

While valuation methods have to deal with a range of difficult-to-value goods, the question of the value of human life is often a lightning rod for attention. Valuing human life is a central concern in particular for policy options that impact on health via differences in mortality. What is actually valued (whether via contingent or revealed-preference methods) is an increase in the risk of death. For example, the wage differences between more or less risky types of otherwise-similar occupations may be compared. On a more practical level, many of the same framing issues discussed above apply to this problem as well and result in a wide range of estimates for the value of a statistical life, for example ranging from one million to eight million dollars in the US (Ashenfelter, 2006). The exact value of a statistical life that is used in an economic evaluation can have a drastic impact on the resulting efficiency of interventions, especially when comparing interventions that save lives with

interventions that do not. Finally, once an estimate is established, questions of equity can also arise, especially between older and younger individuals. CBA takes any life saved as having the same value, while (as we shall see) other methods may value life differently depending on the statistical number of years of life remaining (i.e., based on age).

A second issue that can arise in relation to all valuation methods is that of adaptation to existing conditions (Dolan, 2001), whether negative health states or broader negative health determinants. While in an acute health context, studies try to place value on pure health benefits, healthy public policy often deals with health determinants, and so is interested in valuing these. Studies that either ask the general population for valuations or gather these from market price proxies will tend to reflect the potentially unhealthy social norms of the day. For example, someone who fully depends on his or her car for transportation may not value transportation alternatives much because the car has been naturalized for that person as THE method of getting from A to B; similarly, policies may reflect this and under- or over-price other means of transportation. In response, an argument can be made in favour of asking or studying the choices of those who suffer from a condition to value potential remedies and this could also be extended to health determinants. For example, those who live in very polluted environments may have both the special knowledge of what it is like to live with such a burden and the most to gain from potential policy decisions aimed at remedying it. On the other hand, aside from special knowledge, those who experience a particular condition may have adjusted to it, in addition to the fact that they may have a lesser “ability to pay” than those living in healthier environments. As such, those in the example may be so accustomed to their polluted environment as to place far less value on living with lower levels of pollution. Whether to accept this last argument requires hard thinking about the role of paternalism in public policy.

Finally, people appear to care about their relative position on a continuum of well-being and not just an absolute level (Frank & Sunstein, 2001); in short, this is the idea of “keeping up with the Joneses.” A person may care not only how well off she is, but whether she is as well off as those in some relevant comparison group, such as neighbours, colleagues at work or some other grouping. If this is the case, then equity and social solidarity (both ways of conceptualizing smaller relative differences) may not

figure prominently in valuation methods. Valuation methods that place more value upon absolute benefit outside distributional concerns may undervalue the contributions of increasing inequality.

Cost-benefit analysis (CBA) – Summary

This method transforms all costs and benefits into dollar values, using either revealed preferences or willingness-to-pay to impute price to benefits that do not have clear market prices. It then ranks options according to their Net Present Value (benefits minus costs) or their benefit-to-cost ratio (benefits divided by costs). A Net Present Value (NPV) above \$0 or a benefit-to-cost ratio above 1 means that an option is cost-saving.

Questions to keep in mind to identify ethical implications:

- Are all relevant costs and benefits included in the evaluation?
- Is the evaluation comparing interventions or policies positively or negatively affecting groups of different income levels? If so, is the evaluation comparing **ability-to-pay** instead of **willingness-to-pay**?
- Are there any major biases (scope, budget constraint, framing) that could have affected the results of the evaluation?
- If the value of a statistical life was used, was the figure appropriate and could it have produced different results had it been different?
- Could adaptation to existing conditions have skewed the evaluation results?

C. COST-EFFECTIVENESS ANALYSIS (CEA), OR NON-MONETARY BANG-FOR-THE-BUCK

In the past several decades, cost-effectiveness analysis (CEA) and its offshoots have become important, even dominant, modes of economic evaluation, especially in acute health contexts (Donaldson & Shackley, 2003). Developed in order to counter some of the perceived weaknesses in CBA, CEA focuses on maximizing health, health-related gains or some other appropriate unit rather than purely monetary efficiency (Weinstein & Stason, 1977). CEA has been attractive because it avoids the question of whether benefits can accurately be valued in dollars on the basis of subjective preferences as well as the ethical issues and biases stemming from the use of WTP.

Defining features

Instead of valuing benefits in monetary terms, CEA leaves them in natural units appropriate to the problem being solved. In acute health policy, units can vary from symptom-free days to life-years gained to cases avoided; in broader healthy public policy, they could also include items such as bicycle trips taken, levels of atmospheric pollutants, community centre registrations, the number of people below a poverty threshold or the frequency of food bank use. CEA is a means of including benefits in evaluations that allows for disparate policy options to be easily compared as long as these different options aim to affect the same variable.

Most policy decisions have aims that can be expressed in quantitative measures; CEA isolates these measures and looks at how efficiently they can be positively influenced by a given policy or group of competing policies. For example, a CEA study of several policy options to reduce poverty, an important health determinant, could present its results as the cost per percentage-point reduction in the poverty rate. The most cost-effective policy is then the one for which this cost ratio is lowest. A number of different measures can sometimes be used to look at the same phenomenon – in this example, other measures could include the cost per household whose income rises above the low-income cut-off or the cost per percentage-point reduction in the proportion of households that use food banks. In constructing CEA studies, authors try to focus on measures which best represent the stated policy aims.

Strengths

CEA takes pricing or valuing benefits out of the equation. What matters in CEA is how efficient a given policy option is at securing its desired aim in terms of that aim. The ability to compare the effects of interventions in terms of very specific outcome measures has significant advantages in tractability and relative transparency. Data on many outcome measures for both health and health determinants are either readily available or relatively easy to gather. Such data is also often additionally segmented by gender, age, race and socio-economic status, making it possible for analysts to focus on particular subgroups. Using money as an intermediary to establish the value of various benefits, as in CBA, adds a layer of complexity and ambiguity to economic evaluations. It is possible to ask not only whether a policy option is worthwhile but

whether a benefit is correctly valued. CEA studies remove the second question and allow policy makers to compare the costs of impacting on their desired aims directly. This is done by constructing what are called Incremental Cost-Effectiveness Ratios (ICER) that measure how much a unit gain in the relevant outcome measure would cost over and above a relevant alternative – either the “do nothing” case or another proposed intervention.

Limitations

The concreteness of CEA, however, also gives rise to limitations. Since it lacks a unifying unit of measure, CEA is only able to accommodate one type of effect at a time (this is the textbook version; see cost-consequence analysis [CCA] below for an idea of how multiple measures may figure in practice). If a given policy option positively impacts a number of factors, these benefits have to be looked at separately. The more disparate these factors are, the harder it is to compare them and gauge the overall efficiency or desirability of a particular policy option. This is especially relevant to healthy public policy, where policy aims and effects are often wide ranging. For example, if we are looking at air pollution, then increases in high-quality bicycle infrastructure and more effective, mandatory catalytic converters for cars may both significantly reduce air pollution; however, bicycle infrastructures may also improve health outcomes via greater physical activity.

A second, related limitation is the need for direct comparability. That is, CEA requires that the effects of interventions need to impact on at least some of the same variables for them to be compared. While this is good for gauging the efficiency of various options at achieving a singular or closely related goals, it makes it difficult to look at the allocation of funds between distinct competing uses. There are, however, some more general measures, such as life-years gained (health effect) or CO₂-emission equivalents (environmental effect), that make more broad comparisons possible. Unlike CBA, which looks at all benefits in terms of money and so has a universal standard of comparison, the usefulness of CEA can be limited in debates about which goal, out of a possible range of goals, could be achieved most efficiently with given funds (Rudmik & Drummond, 2013).

Finally, there is a tension between clinical efficacy and population efficacy. Many of the values for outcome measures in CEA evaluations are taken from experimental studies; these, however, may not accurately reflect what happens once an intervention is applied beyond the laboratory or clinical trial. There are many potential dynamics and interacting factors within real-world populations that can change how well an intervention achieves its stated goals.

Ethical implications

CEA does not include the same kind of subjective component in valuing effects, thus lowering the number of ethical issues stemming directly from its application to policy decisions. While CBA is prone to various biases because the valuation of benefits depends on individual opinion, the natural units used in CEA lead to few controversies themselves. Ethical issues arise when we consider which measures are appropriate to a given policy question and how policy benefits are distributed. For example, a CEA of interventions to decrease childhood obesity could rank an intervention focused on after-school fitness programs as very cost effective due to large forecast average changes in program participation and resulting reductions in obesity. It is, however, possible that the gains from this intervention be unequally distributed as the availability, quality and accessibility of after-school fitness programs can be highly dependent on socio-economic status.

Put differently, CEA assumes that a unit of the benefit of interest is worth the same amount no matter how it is gained or to whom it accrues (Schlander, 2010). Examples such as the one above point to the fact that this kind of equal treatment can negatively impact equity. Despite the seemingly clear-cut nature of natural units such as obesity rates, equity considerations can arise in CEA as much as in other evaluation methods.

Cost-effectiveness analysis (CEA) – Summary

This method ranks options that aim at the same goal. It ranks them according to their ratio of cost per X gained, where X can be any natural unit (e.g., life years gained, km of cycle tracks, etc.).

Questions to keep in mind to identify ethical implications:

- Does the evaluation compare interventions or policies which are likely to have positive or negative effects other than the one (variable X) on the basis of which they are compared?
- Does the evaluation compare interventions or policies which are positively or negatively affecting different groups of the population?

D. COST-UTILITY ANALYSIS (CUA), OR THE QALY APPROACH

The limitations imposed by CEA, particularly the lack of comparability across different intervention goals, have led to the development of offshoots that increase flexibility while maintaining the focus on direct measures of health and well-being. The major offshoot of CEA that has gained widespread acceptance in economic evaluation for policy making, particularly in health and health-related disciplines, is cost-utility analysis (CUA). Rather than compare efficiency across single benefits as does CEA, CUA instead analyzes the impact of interventions on a broader measure of quantity and quality of life (Wagstaff, 1991). One of the primary difficulties lies, then, in defining a unit of measure that is at once complex enough to reflect some of the complexity behind the very broad notion of quality of life, but that is also tractable and generalizable across a range of policies and interventions that aim at different health- and wellness-related ends. The measure that has gained the most acceptance is the “Quality-adjusted life year” or QALY (pronounced *kwa-lee*) for short (Weinstein, Torrance, & McGuire, 2009).

Defining features

The QALY is a measure of general health benefit and so can be used even in comparisons of policies that affect a variety of health-related domains. It combines the gain in quality of life over a number of years. Health quality is measured on a scale from 0, signifying death, to 1, signifying full health. Hence, for example, two years at 0.4 quality or one year at 0.8 quality are both worth 0.8 QALYs.

To determine where health states are placed on the 0-to-1 scale, studies are constructed that ask individuals about their health-related preferences through a variety of questionnaires. The results are tabulated and various health states and health improvements are transformed into QALY values. A raspy cough, a broken limb or a life-threatening tumour can each be judged to variously impede quality of life depending on who is judging and in what circumstances. Nevertheless preference-based measures that directly ask people to judge their current or imagined health state are the most prevalent (Brazier, Ratcliffe, Tsuchiya, & Solomon, 2007). The QALY thus has an important subjective component. Who is asked, how they are asked and in what context – all of these factors can impact the final QALY metric.

Once health states are valued in QALYs, interventions can be “costed” by calculating how much each QALY gained would cost under a given intervention – again always in comparison with a relevant alternative. Policy options can be analyzed in terms of their average cost per QALY gained in comparison to a relevant alternative, the “do nothing” case or in comparison to a desired threshold (Simoens, 2009). For example, the UK has for some time adopted a threshold of £20,000-£30,000 (+/- \$35,000-\$50,000) per QALY as acceptable for new health technologies and therapies in acute care (National Institute for Health and Clinical Excellence, 2011) and many other jurisdictions have implemented similar measures.

Strengths

In some respects, CUA is a type of cost-effectiveness analysis (CEA) that shares some of the strengths of cost-benefit analysis (CBA). Like CBA, CUA allows for comparability across a range of interventions that have differing aims. Any policy that impacts health or health determinants can be valued in terms of QALYs gained. Yet, CUA is also more constrained than CBA in that it does not give equal weight to all preferences. Indeed, preferences for goods outside health are completely disregarded (unless they are included in costs).

CUA further differentiates itself by anchoring its valuations of benefit somewhere between pure subjectivity and objectivity. In generating QALY values, subjective estimations of quality modify an objective measure, namely a year of life gained. As such, the QALY is thought to describe something

neither completely objective, such as the natural measures used in CEA, nor completely subjective, such as the dollar values in WTP. Rather, it is something in between. The QALY is an attempt to produce a measure that corresponds to the desires of the population on a closed, 0-to-1 scale that also somehow mirrors an objective health and well-being continuum.

Limitations

CUA was largely developed for use in health economics and economic evaluations of acute care interventions. Although the QALY is a relatively open and broad measure, healthy public policy is much broader than acute care and its impacts on health are often less tangible, spread over a longer timeframe and more indirect. In addition, the main aim of healthy public policy is not always explicitly improving the health of the population. Therefore, some of the intended benefits stemming from healthy public policies might be hard, if not impossible, to capture via CUA.

In addition to this difficulty particular to healthy public policy, several general methodological difficulties stem from the QALY's subjective component. First, a number of factors from the choice of survey method to the way questions are framed to the number of alternatives presented can all influence the values that an individual gives to particular health states (Brazier et al., 2007). Subjective valuations of health are context-dependent to such a degree that there is debate whether they are elicited or, indeed, constructed during a survey (Brazier et al., 2007). Second, evaluations of the same condition can also vary substantially based on general health state as well as a host of socio-economic factors (Dolan, 2001). For example, individuals with a given health condition may either over- or under-value improvements in health compared to values derived from the general population, depending on whether they have grown accustomed to the negative impacts of their condition or strongly desire to be rid of them. Finally, there may be conflict between the objective factors that define health states and individuals' subjective experiences of these same states. That is, while someone may perceive two health states as equivalent from a subjective point of view, one state can have much greater and more debilitating long-term consequences than the other.

Ethical implications

Beyond the methodological limitations just listed, CUA also raises a number of ethical issues. A number of these are related to the equity of intervention outcomes. A central theme concerns the questions: To whom will QALYs accrue in the interventions themselves? Should all health improvements of the same QALY value be treated equally? For some, “a QALY is a QALY,” which means that a QALY is worth the same no matter who receives it – a kind of blanket equality. Others argue that this simple maximization of QALYs could potentially lead to various kinds of discrimination. Below are two examples to illustrate this.

First, some argue that QALY maximization can discriminate against the old as younger individuals may be able to derive greater total benefit from health improvements over the course of their lifetimes (Harris, 2005). For example, providing a health-improving intervention to a 20-year-old saves many more QALYs than providing the same intervention to a 60-year-old simply because the average 20-year-old will live for much longer than the 60-year-old, enjoying, say, 60 rather than 20 years of higher-quality life on average. There have been rejoinders to this view. One type of response has focused on methodological adjustments, suggesting they may be sufficient to counter discrimination claims; changes could include the incorporation of equity weights or the use of alternative questionnaires (Round, 2012). Others suggest a response on the level of alternative ethical principles (for example, the idea that all individuals are entitled to “fair innings” – most broadly, opportunities to attempt to live as one wishes, which are, on average, used up with age) (Tsuchiya, 2000). A somewhat related concern is that regarding those living with disabilities, who some claim are placed in a situation of “double jeopardy” (Harris, 1995): not only do they experience a lower QALY level due to an existing health condition or persistent health determinants, they are also penalized in any further life-extending QALY calculation. For example, persons with chronic obstructive pulmonary disease (COPD) will have, in general, a lower quality of life than people without chronic conditions. If two life-extending interventions are compared, where one benefits people with COPD and one benefits people without a chronic condition, then, assuming the costs are equal, the intervention benefiting people with COPD would have to extend their life much more just to be deemed as efficient as the intervention benefiting

people without a chronic condition, because each year of life gained is worth less QALYs for people with COPD.

Beyond equity concerns about the distribution of quality of life, a final ethical question that has been raised is whether the very notion of adjusting for the quality of life is ethically warranted. Some question this central assumption of CUA on the basis of an ethical principle called the “rule of rescue,” the notion that saving the life of an identifiable person (often thought to be in immediate peril) should trump quality-based cost effectiveness that operates on statistical averages (McKie & Richardson, 2003). Those who support this rule assume that we simply react differently to identifiable others and should implement policies that allow us to act on this kind of reaction. For some proponents, the rule of rescue is ethically valuable rule-following that works, in dire circumstances, to affirm a shared social commitment to human life (MacLean, 1986); others have tried to integrate this type of concern into the ethical framework of maximizing total social good that underlies the QALY approach (McKie & Richardson, 2003). More generally, these kinds of arguments in a milder form point to another kind of equity consideration: for greater attention to more severe conditions.

A different way that equity considerations enter into economic evaluations is through the fact that many interventions are aimed at subgroups, whether demarcated by age, health condition, gender, ethnicity, socio-economic status or some other criterion. Overall, if equity is to be taken into account, then doing so can occur when ascertaining both the inputs of evaluations (the values given to particular effects of policy options) and their outputs (the final distribution of QALYs). Two options for dealing with existing inequities before the presentation of results are to derive QALY values from particular subgroups rather than the general population or to establish separate thresholds for various subgroups (Bobinac, van Exel, Rutten, & Brouwer, 2012). “Segmenting” by subgroup or forming profiles of intended policy beneficiaries can give a truer picture of what policies would mean for those who have the most to gain from them. It can give disadvantaged subgroups a voice (by the proxy of using their QALY valuations). Finally, it could also be a means of acknowledging the background assumption and value that individuals are the best judges of their own welfare, mitigating the risk that the preferences of a general population reference group are generalized and

“imposed” (Birch & Donaldson, 2003) on others. On the other hand, such initial segmenting is rare and difficult in practice. QALY values are generally derived from studies of the general population; as Section “Putting it together: Decision criteria” (pages 14-15) will show, however, attention to subgroups can be deferred until later in the evaluation process.

Cost-utility analysis (CUA) – Summary

This method ranks options according to their ratio of cost per QALY gained (which is a broad measure combining quantity and quality of life) and/or assess whether an option falls below a given efficiency threshold (e.g., \$50,000/QALY).

Questions to keep in mind to identify ethical implications:

- Are some of the evaluated policies or interventions likely to have effects that will not be fully captured by QALYs, either because they are not health-related or they relate to a broader conception of health than QALYs?
- Is the evaluation comparing healthy public policies with acute care interventions? Since healthy public policies may have stated aims other than improving health, their full effects are likely not to be captured by CUA and they may for this reason compare unfavourably to acute care interventions.
- Is the evaluation comparing interventions or policies targeting different age groups? If so, CUA will tend to favour policies or interventions targeting younger people (ageism).
- Is the evaluation comparing some interventions or policies targeting disabled or chronically ill people with interventions targeting people without those conditions? If so, CUA will tend to favour interventions targeting people without those conditions (double jeopardy).

E. COST-CONSEQUENCE ANALYSIS (CCA), OR EXTENDING AND UNITING ANALYSIS

Defining features

With the varying strengths and limitations of different methods, it is not surprising that increasing numbers of evaluations seek to combine elements from different methods to better support policy decision making. Cost-consequence analysis (CCA) is the most prominent means of synthesizing methods and presenting their results. In brief, CCA combines a number of outcome measures in one evaluation, while keeping each one separate. Rather than aiming at a single measure of efficiency, such as a cost per quality-adjusted life year (QALY), CCA commonly presents a table that lists a number of benefit (and cost) measures for each intervention being considered. These measures can include changes in natural units (CEA), changes in QALYs or cost-per-QALY ratios (CUA), particular costs in dollars, particular benefits valued in dollars (CBA), cost-benefit ratios (CBA) and more. An evaluation of a social housing program may include changes to homelessness rates, changes to child poverty rates, changes in QALYs, cost per housing unit, cost per 1000 population and so on.

Strengths

This strategy answers some of the criticisms raised against other methods. For example, while a single-measure CEA is difficult to apply to comparisons of interventions with differing aims, CCA’s inclusion of multiple outcome measures gives policy decision makers greater flexibility in looking at a wide range of interventions. Furthermore, each of CBA, CEA and CUA seeks to provide a single number that represents intervention efficiency and so clearly demarcates costs and effects. While costs are regularly presented in cost tables across methods, the various effects are lumped together and one or a small handful of overall measures of efficiency is often produced as a final result. CCA makes it easier to examine positive and negative outcomes separately. For example, while a QALY measure aggregates beneficial and adverse effects, these can be separated out in a suitably designed CCA. This is especially the case for healthy public policy, where intervention effects are wide-ranging and impact many areas of human life, sometimes in opposing ways. Finally, because CCA presents measures separately, it can offer a greater wealth of information than each alternative method on its own; a CCA evaluation can include changes in QALYs,

changes in natural units, benefits valued in monetary terms, effects of equity calculated separately and so on.

Limitations

At the same time, such disaggregation removes the ability to easily compare interventions from an overall efficiency standpoint. Indeed, the more measures are presented, the harder it can be to get a full picture of an intervention's effects – especially in comparison with others. Separating out measures also means that some intervention effects are presented in natural units and so the same weaknesses of limited comparability that were discussed in the section on CEA also apply in part to CCA. Lastly, since the table presented in a CCA is not a single measure of efficiency, it is necessary to either choose one particular measure (cost per QALY thresholds, cost-benefit ratios, etc.) out of the table as most important – a difficult task – or grapple with synthesizing the various measures into a decision, which brings its own difficulties.

Ethical implications

Disaggregation puts greater onus on decision makers. Decision makers have to be ready to rank the importance of the various efficiency measures presented in a CCA evaluation; with such evaluations, decisions are based on more than ready-made efficiency measures. On one hand, this gives decision makers more flexibility in the ethical principles that guide decisions. Whether they place greater emphasis on economic efficiency, equity, solidarity or other values, those choosing which interventions to pursue can focus on and rank measures to reflect their particular hierarchy of values. For example, a CCA that compares interventions in municipal zoning regulations may include measures of both the health-related as well as the socioeconomic impacts of the various options. Decision makers interested in pure health-related efficiency could focus on CEA or CUA measures, while those interested in equity and poverty reduction could also take into account measures of income distribution or the like.

This kind of ethical flexibility can also be seen as a reflection of a broader conflict between substantive and procedural justice. That is, what is more important: the content of a decision or how it is reached? Methods other than CCA that produce easily comparable efficiency measures side with the latter; transparency is a key value, but it comes at

the cost of a loss of complexity and diversity in decision criteria. In contrast, CCA gives policy makers greater leeway in the conclusions they can draw from an evaluation, but at the cost of procedural clarity. While the recommendations resulting from a CBA or CUA are generally unambiguous and easily reproducible (at least in theory and despite the fact that these recommendations may be taken more or less into account in the actual decision), those resulting from a CCA can differ depending on the values and interests of those receiving the evaluation.

Cost-consequence analysis (CCA) – Summary

This approach combines several methods (CBA, CEA and/or CUA) in one evaluation of efficiency without integrating the results into one comparable unit. It thus presents results in a table, where each option is a row and the results of different methods are in columns.

A question to keep in mind to identify ethical implications:

- What are the most important criteria for this particular decision and how are they reflected in the various methods presented? In other words, on what basis should recommendations be made?

Evaluations and perspective

These final thoughts about the use of CCA point directly to a broader ethical issue that is an important factor across methods of economic evaluation: whose perspective is the correct one to adopt, both in an evaluation and the ensuing decision-making process? No matter which method is used, evaluations can be performed from different perspectives, which can limit both the costs and, in the case of CBA, the effects that are taken into account (Simoens, 2009). Policies can be considered from a societal perspective, from that of a large-scale unit such as a government ministry, agency or regional government unit, from that of a particular site such as a hospital, workplace, or community centre, or from that of an individual intervention beneficiary. Some costs and benefits relevant from one perspective may not be relevant from another. For example, foregone employment earnings may be relevant to an individual who is hospitalized due to a long-term condition resulting from a polluted environment and should also be

taken into account from a societal perspective. These foregone earnings, however, are not of interest to the hospital or broader health administration unit performing the treatment. As a result, from the health unit's point of view acute treatments may seem more cost effective, while socially they may actually be less efficient than other strategies for dealing with pollution, simply because of the costs ignored by the health unit.

The field of healthy public policy is particularly sensitive to such differences in perspective. Costs borne by one ministry or agency often translate into benefits for another unit, possibly also at a different level of government or far removed in time. For example, while the costs of a bike lane are short-term and may be borne purely by a municipal transportation authority, it could be argued that the greater portion of positive health effects end up being reflected in lower health expenditures by provincial and federal health ministries over a much longer time period. There is thus a tension between ensuring completeness in calculating costs and benefits on the one hand and taking on a narrower perspective that may be more relevant to specific decision makers on the other. Some of this tension can be resolved if it is made explicit in the decision-making process.

Within evaluations themselves, the question about perspective becomes one about which perspective should be assumed by those questioned for valuations. There is an important difference between asking people to value things from their own individual point of view or as if they were taking on the societal perspective (Dolan, Olsen, Menzel, & Richardson, 2003). It is unlikely that one perspective is clearly superior and there are good reasons to believe that the perspective undertaken should depend on context, namely on the interventions to which the results of the evaluation will be applied (Dolan et al., 2003), and that, additionally, multiple perspectives may sometimes need to be presented. For example, a decision about how to allocate public funds across several programs impacting aspects of well-being (such as physical fitness or exposure to environmental pollutants) to varying degrees may be more amenable to being valued from a societal perspective. On the other hand, a decision over different means to alleviate the same specific condition, such as a particularly high level of environmental pollutants in one region, may be better served with valuations from an individual point

of view (perhaps even from those to be impacted by the planned intervention).

Given that decision makers may not have input into the design of valuation methods, it is essential that they are aware of the influence the perspective adopted can have on the results of evaluations. It may also be possible for them to actively engage with the results of evaluations by imagining alternative perspectives if necessary. CCA, for example, provides decision makers with significant latitude in choosing the criteria and values that will guide policy choice.

Perspective – Summary

The perspective from which the efficiency of policies or interventions is assessed determines the range of costs and benefits (or effects) that could be considered for inclusion.

A question to keep in mind to identify ethical implications:

- Is the perspective the right one, in the given context, to capture all relevant costs and benefits and other effects of the interventions/policies compared?

Putting it together: Decision criteria

The broadest ethical implication stemming from the use of economic evaluations is that they are limited in the ethical principles they can take on board. The majority of both their proponents and opponents accept that these methods are firmly grounded in utilitarian ethics. This theory defines good and bad as the extent to which the consequences of actions serve to increase or decrease utility, variously defined as pleasure, preference-satisfaction or some other individually based criterion. As such, evaluation methods should be placed in broader decision-making context if one wants to take into account the principles that underlie other ethical theories, notably those in which rights or duties play a role in judging what is good or right (Adler and Posner, 2006).

As mentioned above, there is much to be said for making the ethical considerations and challenges posed by each evaluation method explicit, especially in the face of principles and values beyond efficiency, such as equity and solidarity. Weighing the results and recommendations given by an evaluation against other desirable options may add

another step to the decision-making process, especially given that the former produces tidy values and numbers whereas the latter are notoriously hard to concretely value and quantify.

To deal with this last difficulty, attention may need to be paid to conflicts in the decision-making process. In the simplest case, a CCA with several CEA outcome measures may require decision makers to debate which criterion should be chosen as most representative of the issue at hand, especially if different criteria point to different policy recommendations. Going further, some have suggested that economic evaluations may be made compatible with other ethical principles given certain modifications – for example, “filters” that devalue or exclude those effects that infringe upon the rights of certain groups at the expense of the good of others can mimic rights-based principles (Lowry & Petersen, 2012). Such techniques fall broadly into the category of sensitivity analysis, which looks at what happens to the results of evaluations if certain factors (such as the valuations of specific costs or benefits, the degree of importance we place on future events and so on) are varied.

Another option that is gaining popularity (Marsh, Lanitis, Neasham, Orfanos, & Caro, 2014) is Multi-criteria decision analysis (MCDA), which can be used to systematize and quantify the conflicts between potentially divergent principles. In an MCDA, the authors of evaluations and decision makers have the opportunity to take into account a range of criteria separately; these can include things like economic efficiency and equity, but also access, appropriateness, integration into the health system or community empowerment. Measures of economic efficiency can be used as composites (cost per QALY, NPV) or disaggregated into the values of particular effects. The results are usually presented in a table with criteria along one axis (cost per QALY, ratio of benefit to cost, equity, etc.) and policy options along another. The table can be evaluated qualitatively and a subjective decision made; going further, it is possible to translate the criteria into quantitative scales. Each option is then given a score on each criterion and each criterion is given a weight (e.g., cost per QALY [30%], ratio of benefit to cost [50%] and equity [20%]). This allows decision makers to arrive at a composite score for each option. The process of creating scales and weights allows decision makers to focus on the competition between values and even to involve stakeholders to identifying weights, or the importance of competing

criteria (Baltussen & Niessen, 2006). MCDA allows for a decision procedure that can lead to the choice of a single outcome, while at the same time being procedurally transparent about how ethical value conflicts were resolved.

Finally, as policy effects are calculated according to individual valuations, the resulting perspective reflects what individuals rather than a community as a whole think policy priorities should be. Some worry not only that this is too limiting, but that it also may reflect the values of individuals as consumers rather than as citizens of society with a common interest in the well-being of its members (Mooney, 2009). Efficiency need not necessarily be at odds with the opinions of a more engaged community, but there may be significant benefits to a more deliberative process that involves citizens (Schlander, 2008).

All of the possible biases listed above point to a general ethical challenge for those using the results provided by evaluations: how to acknowledge collective preferences and values, while at the same time being able to question whether they replicate harmful norms or reflect injustices or inconsistencies. Using tools like MCDA or simple explicit consideration of conflicting values is one way that decision makers can navigate this ethical challenge by explicitly acknowledging their own values and biases.

Putting it together: The right method for the job

The outgrowth of new methods of economic evaluation and the refinement of existing methods are good news for the policy community. Despite some of the ethical and methodological challenges described for each method, evaluations are crucial in helping decision makers come to policy choices that both promote public health and get the most out of public funds. Awareness of the strengths and limitations of each method may introduce complications into the decision process, but such awareness is ultimately enabling, especially when value conflicts can be made explicit in the decision process. Individuals armed with more insight into how evaluations are constructed, what biases they can carry and why they lead to the recommendations they do may no longer see them as black boxes but as a more integral part of the policy-making process.

Given the distinct strengths and challenges of each method, there may be situations when one is more appropriate than another or even when two methods can give conflicting results. Indeed, a recent study from the UK (Peters & Anderson, 2012) describes just such a conflict. The authors examined whether mandatory speed zones are a cost-effective means of preventing road injuries. Using both CBA and CUA, they found that the policy recommendation differed depending on the method and background assumptions. In particular, using CUA the authors found that the cost-per-QALY of implementing the speed zones was above the UK's standard threshold for areas with both low and high injury rates. Using CBA, on the other hand, it turned out that although the costs outweighed the benefits in low injury areas, the opposite was true of high injury areas, where benefits outweighed costs and the policy of implementing speed zones would be recommended. The authors noted that CBA can take into account a greater range of benefits, such as the absence of productivity losses, in the calculation of the value of prevented casualties and this may help account for some of the difference in results.

Studies such as the one just described highlight the need to be attentive to all the difficult methodological and ethical issues that can be revealed upon closer examination of economic evaluations. When trying to determine which policy option should be adopted in a given situation, it matters greatly which effects are included, how they are calculated, how they are valued, whose perspective is adopted and how other ethical principles are negotiated. Each method has its own particular biases. While these are unavoidable to some degree, recognizing them can help make decisions clearer and illuminate their rationales. This is especially important in healthy public policy, which is very broad in the scope of policy areas it encompasses and where policies from across different fields are often placed side by side. Attention to the methods of economic evaluation recognizes that every potential intervention and policy choice has to be examined individually. There are no hard and fast rules for efficiency, but much policy work is stronger when efficiency is taken into account with a simultaneously open and critical eye.

Here are some questions stemming from the issues raised in this document that can help guide such an approach to the use of economic evaluations in policy making:

- Is the method of economic evaluation appropriate to the policy question (e.g., is it flexible enough to take into account the important potential effects of the proposed policy or policies)? If not, what other information needs to supplement that provided in the evaluation in the decision-making process?
- Is equity a factor in the particular policy decision? If so, were equity considerations included in the economic evaluation (either in the inputs or at some stage in the decision-making process)? It may be helpful to identify potential biases in different measures and methods and their impact on equity-seeking groups in particular. Finally, to what extent do equity considerations possibly conflict with efficiency considerations, requiring more nuanced decision making?
- What is the appropriate perspective to be taken in a particular policy decision? Would any important costs or benefits be left out or added with a change in perspective? Could the interventions in question benefit from taking multiple perspectives into account (e.g., that of society, an administrative unit, the beneficiaries themselves, etc.)? If so, is it possible to do this within the confines of the existing evaluation or is it necessary to seek out additional information?
- Are there reasons to also focus on community-building and participation in a particular policy decision? If so, is it possible to provide space for community engagement in the policy process that complements the results of evaluations?

References

- Adler, M. D. & Posner, E. A. (2006). *New foundations of cost-benefit analysis*. Cambridge, MA: Harvard University Press.
- Ashenfelter, O. (2006). Measuring the value of a statistical life: Problems and prospects. *The Economic Journal*, 116(510), C10–C23.
- Baltussen, R. & Niessen, L. (2006). Priority setting of health interventions: The need for multi-criteria decision analysis. *Cost Effectiveness and Resource Allocation*, 4(1), 14.
- Birch S. & Donaldson, C. (2003). Valuing the benefits and costs of health care programmes: Where's the "extra" in extra-welfarism? *Social Science & Medicine*, 56, 1121–1133.

- Brazier, J. E., Ratcliffe, J., Tsuchiya, A., & Solomon, J. (2007). *Measuring and valuing health for economic evaluation*. Oxford: Oxford University Press.
- Bobinac, A., van Exel, N., Rutten, F. F., & Brouwer, W. B. (2012). Inquiry into the relationship between equity weights and the value of the QALY. *Value in Health, 15*(8), 1119–1126.
- Brouwer, W. B., Culyer, A. J., van Exel, N., & Rutten, F. F. (2008). Welfarism vs. extra-welfarism. *Journal of Health Economics, 27*(2), 325–338.
- Cookson, R. (2003). Willingness to pay methods in health care: A sceptical view. *Health Economics, 12*, 891–894.
- Dolan, P. (2001). Output measures and valuation in health. In M. Drummond & A. McGuire (Eds.), *Economic evaluation in health care: Merging theory with practice* (pp. 46–67). Oxford: Oxford University Press.
- Dolan, P., Olsen, J. A., Menzel, P., & Richardson, J. (2003). An inquiry into the different perspectives that can be used when eliciting preferences in health. *Health Economics, 12*, 545–551.
- Donaldson, C. & Shackley, P. (2003). Willingness to pay for health care. In A. Scott, A. Maynard, & R. Elliott (Eds.), *Advances in health economics* (pp. 1–24). Chichester: John Wiley & Sons, Ltd.
- Frank, R. H. & Sunstein, C. R. (2001). Cost-benefit analysis and relative position. *University of Chicago Law Review, 68*, 323.
- Gold, M., Siegel, J., Russell, L., & Weinstein, M. C. (1996). *Cost-effectiveness in health and medicine*. New York: Oxford University Press.
- Harris, J. (1995). Double jeopardy and the veil of ignorance—a reply. *Journal of Medical Ethics, 21*(3), 151–157.
- Harris, J. (2005). It's not NICE to discriminate. *Journal of Medical Ethics, 31*(7), 373–375.
- Hausman, D. M. (2011). *Why satisfy preferences?* (No. 1124). Papers on economics and evolution.
- Jefferson, T., Demicheli, V., & Mufrod, M. (2000). *Elementary economic evaluation in health care*. London: BMJ Books.
- Lowry, R. & Peterson, M. (2012). Cost-benefit analysis and non-utilitarian ethics. *Politics, Philosophy & Economics, 11*(3), 258–279.
- MacLean, D. (1986). Social values and the distribution of risk. In D. MacLean (Ed.), *Values at risk*. New Jersey: Rowman and Allanheld.
- Marsh, K., Lanitis, T., Neasham, D., Orfanos, P., & Caro, J. (2014). Assessing the value of healthcare interventions using multi-criteria decision analysis: A review of the literature. *PharmacoEconomics, 32*(4), 345–365.
- McKie, J. & Richardson, J. (2003). The rule of rescue. *Social Science & Medicine, 56*(12), 2407–2419.
- Mooney, G. (2009). *Challenging health economics*. Oxford: Oxford University Press.
- National Institute for Health and Clinical Excellence (NICE). (2011). *Discounting of health benefits in special circumstances*. London: NICE.
- Peters, J. L. & Anderson, R. (2012). The cost-effectiveness of 20 mph zones for the prevention of injuries. *Journal of Public Health, 35*(1), 40–48.
- Round, J. (2012). Is a QALY still a QALY at the end of life? *Journal of Health Economics, 31*(3), 521–527.
- Rudmik, L. & Drummond, M. (2013). Health economic evaluation: Important principles and methodology. *The Laryngoscope, 123*(6), 1341–1347.
- Schlander, M. (2010). Measures of efficiency in healthcare: QALMs about QALYs?. *Zeitschrift für Evidenz, Fortbildung und Qualität im Gesundheitswesen, 104*(3), 214–226.
- Simoens, S. (2009). Health economic assessment: A methodological primer. *International Journal of Environmental Research and Public Health, 6*(12), 2950–2966.

Sunstein, C. R. (2013). The real world of cost-benefit analysis: Thirty-six questions (and almost as many answers). *Columbia Law Review*, 114, 167–211.

Tsuchiya, A. (2000). QALYs and ageism: Philosophical theories and age weighting. *Health Economics*, 9(1), 57–68.

Wagstaff, A. (1991). QALYs and the equity–efficiency trade-off. *Journal of Health Economics*, 10, 21–41.

Wegner, G. & Pascual, U. (2011). Cost-benefit analysis in the context of ecosystem services for human well-being: A multidisciplinary critique. *Global Environmental Change*, 21(2), 492–504.

Weinstein, M. C. & Stason, W. B. (1977). Foundation of cost–effectiveness analysis for health and medical practices. *New England Journal of Medicine*, 296, 716–721.

Weinstein, M. C., Torrance, G., & McGuire, A. (2009). QALYs: the basics. *Value in Health*, 12(s1), S5–S9.

November 2014

Author: Michal Rozworski, Independent Researcher

Editing: Olivier Bellefleur, National Collaborating Centre for Healthy Public Policy

SUGGESTED CITATION

Rozworski, M. (2014). *Methods of economic evaluation: What are the ethical implications for healthy public policy?* Montréal, Québec: National Collaborating Centre for Healthy Public Policy

ACKNOWLEDGMENTS

The NCCHPP would like to thank the following reviewers:

Astrid Brousselle, Associate Professor, Département des sciences de la santé communautaire, Centre de recherche de l'hôpital Charles-LeMoine, Université de Sherbrooke, Canada Research Chair in Evaluation and Health Care System Improvement

Claire Gram, Population Health Policy Consultant, Vancouver Coastal Health

Gabriel Tremblay, Senior Health Economist, Adelphi Values

The National Collaborating Centre for Healthy Public Policy (NCCHPP) seeks to increase the expertise of public health actors across Canada in healthy public policy through the development, sharing and use of knowledge. The NCCHPP is one of six centres financed by the Public Health Agency of Canada. The six centres form a network across Canada, each hosted by a different institution and each focusing on a specific topic linked to public health. The National Collaborating Centre for Healthy Public Policy is hosted by the Institut national de santé publique du Québec (INSPQ), a leading centre in public health in Canada.

Production of this document has been made possible through a financial contribution from the Public Health Agency of Canada through funding for the National Collaborating Centre for Healthy Public Policy (NCCHPP). The views expressed herein do not necessarily represent the views of the Public Health Agency of Canada.

Publication N°: 2743

This document is available in its entirety in electronic format (PDF) on the Institut national de santé publique du Québec website at: www.inspq.qc.ca/english and on the National Collaborating Centre for Healthy Public Policy website at: www.ncchpp.ca.

La version française est disponible sur les sites Web du Centre de collaboration nationale sur les politiques publiques et la santé (CCNPPS) au www.ccnpps.ca et de l'Institut national de santé publique du Québec au www.inspq.qc.ca.

Reproductions for private study or research purposes are authorized by virtue of Article 29 of the Copyright Act. Any other use must be authorized by the Government of Québec, which holds the exclusive intellectual property rights for this document. Authorization may be obtained by submitting a request to the central clearing house of the Service de la gestion des droits d'auteur of Les Publications du Québec, using the online form at <http://www.droitauteur.gouv.qc.ca/en/autorisation.php> or by sending an e-mail to droit.auteur@cspq.gouv.qc.ca.

Information contained in the document may be cited provided that the source is mentioned.

LEGAL DEPOSIT – 4th QUARTER 2020
BIBLIOTHÈQUE ET ARCHIVES NATIONALES DU QUÉBEC
LIBRARY AND ARCHIVES CANADA
ISBN: 978-2-550-87863-6 (FRENCH PDF)
ISBN: 978-2-550-87864-3 (PDF)

©Gouvernement du Québec (2020)

