Cost-Effectiveness Analysis

5

Jeroen Luyten and Catherine Henderson

Abstract

Economic evaluation of healthcare programs seeks to compare treatments and preventive measures in terms of their efficiency, that is, their ability to generate health and well-being relative to the costs incurred. This chapter provides an introduction to one particular but widely used evaluation technique: cost-effectiveness analysis (CEA). We present the main conceptual elements of a CEA, measurement techniques that are used, and the challenges and limitations, and we discuss the final interpretation of results within the context of the mental health field.

Key Points Summary

- Costs
- Outcomes
- Discounting
- Cost-effectiveness ratios
- Net benefits
- Uncertainty

J. Luyten (✉)
Department of Social Policy, London School of Economics and Political Science, London, UK

Leuven Institute for Healthcare Policy, KULeuven, Kapucijnenvoer 35, 3000 Leuven, Belgium
e-mail: J.Luyten@lse.ac.uk;
Jeroen.luyten@kuleuven.be

C. Henderson
Personal Social Services Research Unit (PSSRU),
London School of Economics and Political Science,
London, UK
e-mail: C.Henderson@lse.ac.uk

5.1 Introduction

Purchasers and planners of mental health services need to make investments that achieve the best results for their patients using available resources. Some guidance for deciding how to make these investment choices is required. For instance, how should a decision maker determine how to divide funds between different treatments for depression (such as cognitive-behavioral therapy, antidepressant medication, and psychotherapy)? The decision-maker will naturally want to choose among the most effective treatments, but there is also an unavoidable economic aspect to this choice (see Chap. 10). The resources necessary to make treatments available are, by definition, finite, in terms of not only funding but also health personnel, treatment spaces, and infrastructure. A central concern for decision makers who have to manage these resources is

therefore to provide a mix of treatments that maximize desired mental health outcomes for patients. Or, in economic terms, to allocate resources in a way that minimizes opportunity costs (see Chap. 1), the value of the "next best alternative use" of a resource that is not chosen and is consequently lost forever. The opportunity cost of providing one treatment for depression is the loss of another treatment that could have been provided instead, at the expense of the potential benefits to patients of that other treatment.

Allocating resources based on minimizing opportunity costs is complex and requires extensive counterfactual information. Economic evaluation has been developed as a standardized and evidence-based technique to facilitate decision making based on opportunity costs [1–3]. It has become increasingly influential in health policy making [4, 5], often with a formal role in many policy contexts, most notably in health insurance coverage decisions (e.g., in the United Kingdom, France, Germany, Belgium – but not in the United States). An economic evaluation compares the costs and outcomes that are linked to at least two interventions, one of which is often the current practice of usual care. Different forms of economic evaluation exist (see Box 5.1). They all have a common approach to costs (see Sect. 5.2) but differ in their assessment of consequences. Depending on the level at which resources need to be allocated and opportunity costs need to be assessed – broad or narrow – one particular type of analysis will be more appropriate than the others. Cost-benefit analysis (CBA) (see Chap. 4) is the broadest form. It assesses consequences in monetary terms so that the return on investment from spending a sum of money in one program can be compared with investing that same sum in any other program – within the health sphere but also beyond, for example, by investing these resources in public infrastructure. Cost-utility analysis (CUA) is limited to comparisons within the health domain [6]. Consequences are expressed in generic health units that compose the effects of a condition on both mortality and morbidity, such as quality-adjusted life years (QALYs), disability-adjusted life years, or healthy year equivalents. This enables opportu-

nity costs of health programs to be assessed in terms of the health units forgone by not investing these resources in competing health programs. Cost-effectiveness analysis (CEA) is a narrower form of assessing opportunity costs in which the assessed consequences are more specific and limited to a particular field of healthcare, mostly one specific disease area. In this chapter, we outline the main elements of cost-effectiveness studies and their interpretation.

5.2   Main Elements of Cost-Effectiveness Studies

5.2.1   Costs

Costs are a function of the volume of the resources consumed when making an intervention available, multiplied by their respective unit cost (see Chaps. 2 and 11). Distinguishing quantities from unit costs is important because it allows the reliability and relevance of the valuations made to be assessed. It also allows assessment of the transferability of results from the original study to other contexts (e.g., countries or times).

5.2.1.1 Resource Use
Typical resources that are used by providing a program are consumables (e.g., pharmaceuticals), labor (e.g., nursing, caregiving), capital (buildings, devices, equipment), and overhead costs (e.g. electricity, management) (see Chaps. 2 and 11). In the domain of mental health, resources consumed outside of the health sector could also be relevant: costs associated with criminal justice, provision of special housing, social care, and additional costs falling on schools because of special educational needs [7] (see Chap. 2). Box 5.2 gives a possible classification of different types of costs that should be considered in a CEA (see Chap. 2).

Overhead costs, such as those of common equipment, personnel, or facilities, can be attributed to individual interventions by relating the proportion of resources used by the intervention relative to the total potential use of the resources, for instance, the number of hours a facility can be

Box 5.1 Different Types of Full Economic Evaluations

- *Cost-minimization analysis* compares the costs of different programs that broadly lead to the same result. Because uncertainty always exists around costs and expected outcomes, in reality, the effectiveness of two programs can rarely be assumed as being equal.
- *Cost-effectiveness analysis* compares the costs and health effects of two or more interventions. Health outcomes in a cost-effectiveness analysis are expressed in terms of specific clinical or other "natural" end points that are measurable and that can be considered important within a particular health domain.
- *Cost-utility analysis* is a broader form of economic evaluation in which health outcomes are both measured and valued. Outcomes are translated into a generic measure of overall health. Several generic outcome measures are available, but the most widely used are the QALY (a measure of health) and the disability-adjusted life year (a measure of illness, mostly used in low- and middle-income contexts).
- *Cost-benefit analysis* is the broadest form of economic evaluation. It assesses health consequences in the most common metric used to assess value: money. Expressing the health effects of an intervention in monetary terms and comparing them with the costs associated with that intervention allows decision makers to judge the return on investment of a program, that is, how much net value an intervention offers. This estimate can consequently be compared with other interventions for which the benefits can also be expressed in monetary terms, both within healthcare and beyond.

(continued)

(continued)

- *Cost-consequences analysis* presents a range of outcomes (measured in natural units) alongside the costs of alternative programs, without defining any one outcome as primary.

Box 5.2 Types of Costs

- **Direct costs** often represent the healthcare resources used by providing a program: doctors' hours, medications, hospital beds, overhead costs of running facilities, capital costs of buildings, training, or equipment. In mental health the costs of other forms of care (e.g., social care) also could be considered direct costs.
- **Indirect costs** are the opportunity costs of patients and caregivers in terms of time lost through ill health, undergoing treatment, or providing unpaid care. These costs mainly represent productivity losses due to an inability to work because of illness, but they could also include disrupted domestic, educational, social, and leisure activities.
- **Patient costs** are those costs borne by patients and their families, such as transport costs, user charges, and time lost. They can be substantial but are often not considered in analyses from a payer perspective.
- **Future costs** are often split between future costs that are directly related to the disease or the intervention (e.g., a mental health problem that gives a higher risk of developing diabetes), and those costs that are unrelated (e.g., increased life expectancy leading to higher pension costs).
- **Intangible costs** are the psychological "costs" of pain and suffering that patients experience during an episode of illness or while undergoing the treatment. These are obviously difficult to quantify.

used to provide a treatment as a proportion of the total hours the facility is available for medical use.

Which cost categories should be considered in a CEA depends on the perspective from which the analysis is undertaken (see Chaps. 1 and 2): that of the patient, the employer, the hospital, the healthcare payer, or society. If a healthcare payer perspective (e.g., national health insurance) is adopted, only those costs that are incurred by the payer should be considered. These primarily include the direct costs of providing the program (other costs predominantly falling on other parties). Analogously, from a patient, employer, or hospital perspective, only the costs borne by those groups are relevant. If a societal perspective is adopted, however, all costs borne by the whole of society should be considered. The benefit of the societal perspective is that it does not neglect any economically relevant costs. A disadvantage is that it does not consider how these costs are distributed among the various affected parties.

The choice of costing perspective can have a substantial effect on the estimated costs of an intervention. This is especially true in the field of mental health. For instance, prevention of depression is much more cost-effective from a societal perspective than from a payer perspective, as the bulk of the cost burden is indirect, attributable to an inability to work rather than to costs associated with healthcare treatment (see Chap. 25). For instance, an English study estimated that 90% of the societal cost of depression was due to unemployment and absenteeism from work [8]. From a patient or employer perspective, it is possible for the (tangible) costs of depression to be lower than the costs for society or the healthcare payer. For instance, if disability payments (state benefits or social insurance payments) sufficiently compensate patients for loss of income, or if employers quickly find replacement employees, then the cost to patients and employers could be minimal (see Chap. 29). Consequently, from a financial perspective, prevention of depression is more or less attractive depending on whose costs are considered.

5.2.1.2 Unit Costs

The resources consumed as part of a treatment program must be valued. Unit costs are mainly understood as prices or charges and can be accessed via national price lists or data on purchasing prices from institutions (e.g., hospitals). The level of detail required depends on the importance of the particular item, the scope of the study, and the time and resources available for the analysis. We can illustrate this point by considering the unit cost of hospital stays. It is less precise but more convenient to use a general per-day cost calculated on the basis of the total cost of the hospital or one of its departments. More precise estimates take into account the particular characteristics of the admission and the treatment, down to the specific resource use of an individual patient (micro-costing) (see Chaps. 2 and 11).

However, some resources (e.g., volunteer time from caregivers) do not have market prices. This obviously does not mean that they are without value, and a costing method that does not account for this use of nontradeable resources would underestimate the opportunity cost of a program. In those cases, a value may be imputed to approximate the value of the resource should there be a market in which the resource could be bought (see Chap. 15). For instance, caregiver time can be valued at average market wage or at hourly wages for overtime (see Chap. 17). Several valuing techniques exist to put a monetary figure on nonmarket resources, most notably contingent valuation (willingness-to-pay or willingness-to-accept studies) (see Chap. 4).

The unit cost that should be used also depends on the costing perspective that is adopted. From a payer, patient, or employer perspective, the market price is often the price actually paid, and it consequently reflects the actual economic loss incurred by the payer, patient, or employer. From a societal point of view, arriving at a valuation can be more complex. What matters here is the change (i.e., the loss) in available economic resources within a country. Market prices of used resources can be misleading in terms of reflecting the true social cost of using these resources.

Hospital charges may reflect cross-subsidization across departments and could artificially inflate or deflate the economic loss incurred by providing one type of treatment. Drug prices often reflect monopoly profits and, depending on the recipient and usage of these profits (e.g., domestic or foreign pharmaceutical companies that either reinvest profits or not), the social loss will be larger or smaller. Moreover – and this is also relevant to payer or patient perspectives – unit costs can become variable (see also Sect. 5.2.5 on marginal cost-effectiveness ratios). Being subject to supply-and-demand dynamics, the prices (and opportunity costs) of particular resources can increase or decrease as a function of the quantity needed. For instance, the value of one unit of nursing time depends on alternative deployment possibilities, and this value will likely be higher when more time is needed. As an example, in the initial stages of an epidemic, spare capacity in a nursing service can be used, but gradually higher opportunity costs will be incurred as more nursing time is taken from other, more productive activities [9]. A fixed unit cost (e.g., an hourly wage) does not reflect such dynamics. These issues of finding appropriate unit costs highlight difficulties in assessing the "true" societal cost of diseases. Obviously, social opportunity costs cannot be a requirement for every single CEA, and the label "societal perspective" is often used for an analysis that just uses indirect costs in addition to direct costs, all valued at listed prices. But it is important to highlight that the value attributed to resources must in some cases be treated with caution, especially for resources that are used in large quantities and for which there are reasons to believe that official prices do not reflect the value of alternative deployments.

5.2.2 Outcomes

A focus on costs only (a cost analysis) might indicate that mental health programs can lead to cost savings (when a sufficiently long time horizon is considered). If a decision maker's only concern is to contain or reduce costs, this information may be sufficient to identify the preferred program. Full economic evaluations, aiming to inform the decision maker of the value received per amount invested in an intervention, also take into account the benefits received for the costs incurred. Estimating the net health effect of an intervention – the denominator of cost-effectiveness – consists of two separate tasks: defining relevant outcomes and measuring them.

5.2.2.1 Defining Outcomes

Ideally, a single and unambiguous outcome (an event, a biological marker, a disease stage, reduction of a specified risk factor) needs to be achieved so that the alternatives being evaluated can be compared in terms of their achievement. This outcome measure needs to be observable, relatively easy to measure, and meaningful in the particular disease context. A "final" outcome, such as depression-free days, might be useful in some study contexts; in others, however, a measure that can be linked to a final outcome (an "intermediate" outcome) may be more relevant or feasible. For example, detecting suicidal ideation could lead to the prevention of death by suicide (the final outcome). Drummond et al. [2] recommend that analysts should explain why the intermediate end point has value or clinical relevance in its own right, be confident that the link between the intermediate and final health outcomes has been adequately established by previous research, or ensure that any uncertainty surrounding that link is adequately characterized in the study.

CEA has a narrower range of applications than a CUA. Nonetheless, CEA may be the natural choice in certain circumstances. It may be that clinicians are very interested in the effect of a treatment on a particular clinical outcome; a CEA based on that outcome might produce evidence that clinicians see as more relevant than a CUA. Clinicians' perceptions of the relevance of the outcome could influence their decision to implement that treatment. In addition, generic preference-based measures (from which utilities are derived) may not perform equally well across all mental health conditions when measuring clinically relevant change. For instance, the evidence is mixed on the validity and responsiveness of the

EuroQol five-dimension questionnaire and SF-6D in measuring the effects of schizophrenia and psychotic disorders [10–13] (see Chaps. 3 and 6). Thus limitations may exist in assessing utility on the basis of these measures in these populations. A CEA based on condition-specific measures of quality of life or symptom rating scales might be considered here in order to adequately capture changes brought about by the intervention [12, 14]. One approach in this circumstance would be to carry out both a CEA and a CUA within one study [compare with refs. 15, 16].

Table 5.1 provides an overview of outcome measures that have been used in economic evaluations within some clinical areas of mental health.

It needs to be said, however, that pinning down the most relevant end point can be complex for many diseases, and there will often be disagreement on the best measure to judge the effectiveness of an intervention. Mental health conditions are often multidimensional. A solution to this issue – at least for researchers – is to expand a CEA into a *cost-consequences analysis*. This is a variant of CEA whereby, instead of defining one single outcome, a range of output measures is presented to decision makers, without judging which measure is the more relevant one.

5.2.2.2 Study Designs

The effects of treatments (and also costs) are likely to differ between individuals. Moreover, different individuals may undergo different treatment regimens, experience the course of a disease differently, and respond differently to treatment. The quality of a CEA is often judged based on the quality of its underlying effectiveness assessment and the extent to which it manages to account for patient heterogeneity. Different study designs have different weaknesses.

In a randomized controlled trial (RCT) with adequate power and appropriate follow-up

Table 5.1 Examples of outcome measures used in cost-effectiveness analyses in mental health

Clinical area	Outcome measure	Studies using the measure
Schizophrenia	Clinical Global Impressions scale (CGI) [17]	King et al. [18]
	Positive and Negative Syndrome Scale for Schizophrenia (PANSS) [19]	Priebe et al. [20]
	Investigators Assessment Questionnaire (IAQ) [21]	King et al. [18]
	Global Assessment of Functioning [22]	Hastrup et al. [23]
Depression	Beck Depression Inventory [24, 25]	Hollinghurst et al. [16]; Kuyken et al. [26]; Maljanen et al. [27]
	Hospital Anxiety Depression Scale (HADS) [28]	Romeo et al. [29]
	Cornell Scale for Depression in Dementia (CSDD) [30]	Banerjee et al. [31]
	General Health Questionnaire, 28-item version (GHQ-28) [32]	Woods et al. [33]
	Time to relapse (using the Structured Clinical Interview for DSM Disorders) [34]	Kuyken et al. [26]; Kuyken et al. [35]
Dementia	The Neuropsychiatric Inventory (NPI) [36]	D'Amico et al. [37]
	Cohen-Mansfield Agitation Inventory (CMAI) [38]	Chenoweth et al. [39]
	Quality of Life in Alzheimer's Disease measure (QOL-AD) [40]	Woods et al. [33]; D'Amico et al. [41]; Orgeta et al. [42]
	The Alzheimer's Disease Assessment Scale – Cognition (ADAS-Cog) [43]	D'Amico et al. [41]; Orgeta et al. [42]
Substance abuse	Addiction Severity Index (ASI) [44]	McLellan et al. [45]
	Days of abstinence (using the Global Appraisal of Individual Needs [GAIN]) [46]	McCollister et al. [47]; McLellan et al. [45]
	Longest duration of abstinence (based on laboratory sampling)	Olmstead et al. [48]
Suicide	Beck Scale for Suicide Ideation [49]	van Spijker et al. [50]

duration, health effects can be recorded on an individual patient basis and can later be causally attributed to the treatment. Adequate randomization across treatment and control groups ensures that other characteristics that might cause differences in effectiveness (confounders) are equally prevalent in both groups. An RCT can also record resource use by individual patients, after which average costs and effects can be calculated [51]. However, RCTs can be costly to carry out and take a long time to complete. Further, depending on the nature of the trial, the outcomes are the products of treatment regimens conducted in ideal circumstances to assess whether the treatment *can* work. Such trials are unlikely to be fully representative of the costs and outcomes of day-to-day clinical practice. This difference between efficacy and effectiveness needs to be considered carefully when relying on RCTs for CEA. More pragmatic RCT designs test the effectiveness of a treatment within routine clinical practice settings [52], for instance, by avoiding the imposition of rigid inclusion and exclusion criteria to reflect the real-world patient population who would receive the treatment. An example of a pragmatic trial in the mental health field is a pharmacological trial comparing classes of antipsychotic medications for people with chronic schizophrenia [53–55]. It is also important to make sure that the control group actually represents the "do nothing" option that is implemented in a particular setting, and that the additional benefit of a program is not overestimated or underestimated by comparing it to an irrelevant alternative (see also Sect. 5.2.4). Alternative study designs such as observational cohort studies have been advocated on the grounds that they can be carried out without the strictures imposed by randomization that may limit the generalizability of findings [56]. Observational studies provide information on the effectiveness of interventions and are important bases for CEAs. Attributing the outcomes of controlled but not randomized studies to the intervention of interest can, however, be affected by confounding due to a lack of random assignment of patients to treatment and control groups. Selection bias has traditionally been a weakness of these designs, but

alternative approaches involving statistical methods for creating "synthetic control groups" are coming into use [56].

When experimental studies are not feasible because of financial, practical, or sometimes even ethical concerns, modeling can be an alternative basis for economic evaluation analyses [57] ((see Chap. 7). Models can be used to project the evolution of a condition in a population, based on a combination of available insights obtained from published estimates. A model allows a simplified depiction of possible consequences resulting from different treatment choices or events. Two popular techniques are decision trees [58], generally used for acute events, and Markov models [59], mostly used to synthesize events that require a longer time frame, as is often the case in mental health.

5.2.3 Discounting

It is important that the time frame considered in an evaluation is long enough so that it captures all relevant aspects of the alternatives under evaluation. But costs and outcomes may occur on separate time scales. In economic evaluation, it is standard practice to revalue costs and effects, depending on whether they occur at more distant or more proximate moments in time. The "present value" (PV) represents the contemporary value of a cost or outcome X occurring n years from now, depreciating at a yearly discount rate of r:

$$PV(X) = \frac{X}{(1+r)^n}$$

Discounting can be contentious, especially when applied to health outcomes [60]. When it comes to costs, there are convincing reasons to account for time preference. First, the future is uncertain; various catastrophic events might occur that would invalidate projected future costs. Second, a sum of money at our disposal now can be invested and generate a larger amount later. If we were to pay a cost in the future but need to account for it now, we would only need to pay a fraction of it. Third, if people are wealthier in the future than they are now, a sum of money at

current prices would represent a smaller proportion of the funds available later. Fourth, as additional units of income will at some point lead to decreasing marginal levels of utility, the relative sacrifice of that cost (its opportunity cost in terms of consumption of the forgone alternative) will likely be lower in the future. And fifth, people tend to have an innate pure time preference (or bias) for the present over the future. We prefer to enjoy life now and to pay later.

Whether these arguments for discounting costs also hold for discounting health outcomes is less clear. Health seems to be a normal (or even luxury) good, rather than one of necessity: as our income grows, we are likely to attribute higher values to extra health gains. Moreover, we cannot invest health over time like we can with money. And a pure time preference may be less pronounced for health than for costs (becoming sick now or in 10 years vs. paying a cost now or in 10 years). On the other hand, health, as with money, arguably has decreasing marginal utility over time: an 85th year may be less valuable than a 65th one. Whether the extra gains from years lived in more prosperous times outweigh the decreased marginal utility of greater longevity is an open question. Also, not applying a discount rate for health gains while discounting costs can create problems of inconsistency and could lead to counterintuitive results. Every program seems better the longer it is postponed into the future (as costs would be discounted but health effects would not). And some interventions (e.g., disease eradication programs) have benefits that last indefinitely. Refraining from discounting these benefits would lead us to overinvest scarce resources in such programs. Last, plenty of empirical evidence shows that people de facto discount health gains in practice, for example, smokers who prefer short-term pleasure to long-term health.

Discounting can have a substantial effect when interventions aim to generate lasting and long-term effects, which is often the case in mental health programs. Most guidelines propose using a well-defined discount rate for costs and often a smaller one for health effects, but recommend presenting results with different rates as well [e.g., 61, 62].

5.2.4 Analytical Methods

5.2.4.1 Cost-Effectiveness Ratios

Combining (discounted) costs and effects, we can derive an *average* cost-effectiveness ratio (ACER), a *marginal* cost-effectiveness ratio (MCER), and – most often reported – an *incremental* cost-effectiveness ratio (ICER).

The ACER expresses the total costs of an intervention per achieved health outcome, as compared with a baseline situation, which in many cases would be the current situation (usual care):

$$ACER_A = \frac{\text{Cost intervention } A}{\text{Effects intervention } A}$$

The MCER expresses the changes in cost and effect within one program when it is expanded in scale (e.g., an education program that is rolled out in two regions instead of one). If the size of program is flexible, the MCER can give a useful indication of the economies of scale that can occur, which is informative in finding the optimal level of program provision:[1]

$$MCER_{A_Q} = \frac{\begin{array}{c}\text{Cost intervention } A \text{ } at \text{ scale } Q \\ +1 - \text{Cost intervention } A \text{ } at \text{ scale } Q\end{array}}{\begin{array}{c}\text{Effect intervention } A \text{ } at \text{ scale } Q \\ +1 - \text{Effect intervention } A \text{ } at \text{ scale } Q\end{array}}$$

The most common form of expressing the results of an economic evaluation is the ICER, comparing the costs and effects of the two most relevant interventions under evaluation. The ICER gives an indication of the extra (or incremental) cost of one program for the extra effect it generates over another:[2]

$$ICER_{AvsB} = \frac{\text{Costs intervention } A - \text{Costs intervention } B}{\text{Effects intervention } A - \text{Effects intervention } B}$$

---

[1] Strictly speaking, the "marginal" value of a variable is its rate of change (first derivative) with respect to quantity. This is equivalent to the formula provided if Q is sufficiently large.

[2] Note that when we are evaluating only one intervention and the comparator intervention is the "do nothing" scenario, the ICER is the same as the ACER.
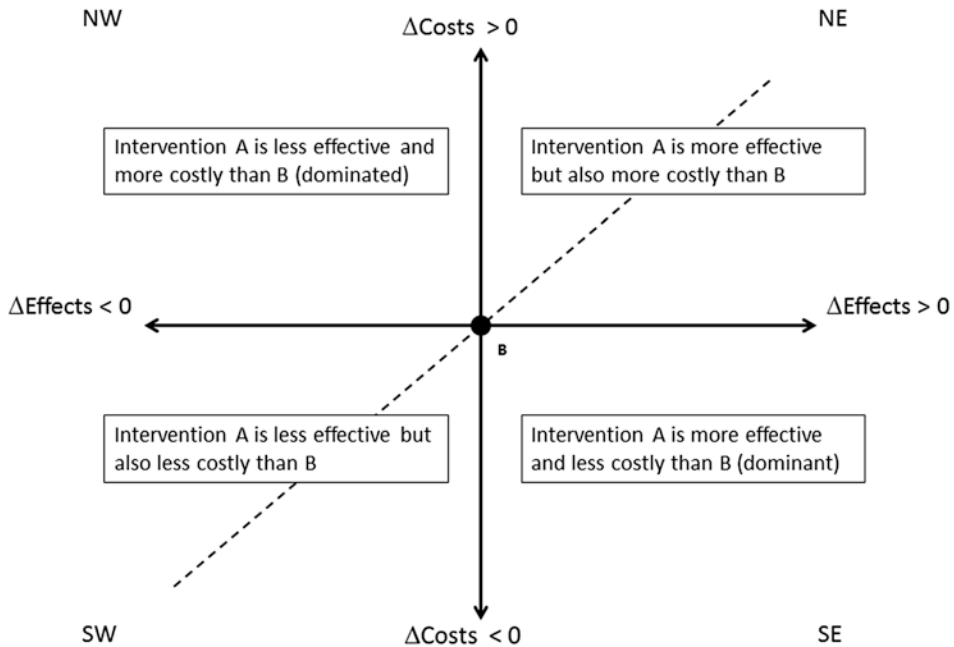
Fig. 5.1   The cost-effectiveness plane

ACERs, MCERs, and ICERs can be represented on a "cost-effectiveness plane," represented in Fig. 5.1. The plane has four quadrants, corresponding to the four main possible outcomes of CEA. An intervention can be more costly and lead to fewer health gains than another one (quadrant NW); in this case the "do nothing" strategy is represented in the origin O. If so, the new intervention is said to be "dominated" by the other one (so we should do nothing). Conversely, if the new intervention is less expensive but leads to better outcomes, it is said to be "dominant" over the other strategy (quadrant SE). More difficult questions arise when one intervention is both more expensive and more effective than the other (quadrant NE). In that case we need to judge whether paying more for better outcomes is "worth it." Similarly, if an intervention is less costly but also less effective than the alternative, are the cost-savings worth the health losses (quadrant SW)? The question "is it worth it?" can only be answered when we have an estimate of the maximum monetary value of the health effect in question, for example, a societal willingness to pay per health effect (represented in Fig. 5.1 by the dashed line through the origin; see also Sect. 5.3).

Analysts need to be cautious when interpreting cost-effectiveness ratios (CERs) with regard to whether the ratio represents the most relevant information. Several implementable strategies may be available, and not all ICERs are ultimately meaningful. Figure 5.2 illustrates the different types of CERs and how to exclude irrelevant ones. The slopes of all lines connecting the points are all CERs. The slopes of the lines starting from the origin (the "do nothing" strategy) represent the ACER of each strategy, indicating how much the average gain in effect would cost in each strategy. A′ is intervention A scaled up by one unit. The dotted line connecting intervention A and A′ represents the MCER of A′ versus A, and doing this for the entire range of possible output levels provides information about the optimal level of program provision for A. In this case, the slope between A and A′ is smaller than the one between O and A, indicating economies of scale: the same health effect can be offered at a lower unit cost, for instance, because of the fixed costs of starting up the program. When several programs are available (A, A′, B, C, D, and X), the analyst must plot the costs and outcomes on the cost-effectiveness plane and

Fig. 5.2 Cost-effectiveness ratios of different strategies
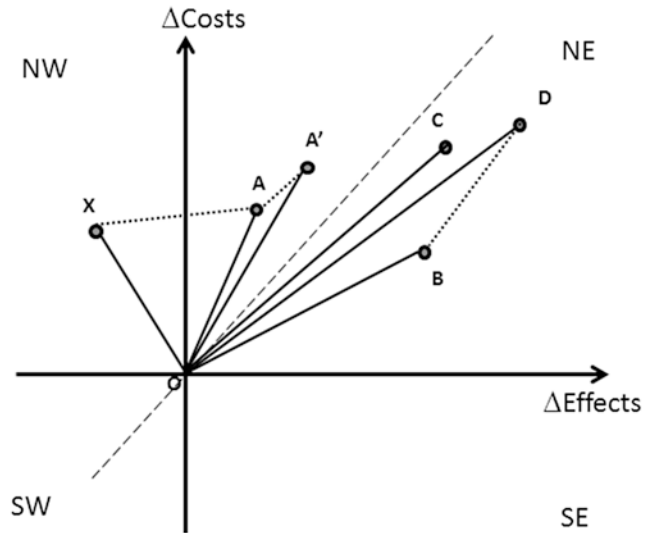
eliminate those strategies that are dominated. In this example, interventions A, A′, X, and C are all dominated by B and D. A, A′, and X are "strictly dominated" by intervention B (i.e., B leads to better health outcomes at a lower cost). C is dominated by extension ("extended dominance") because the combination of strategies B and D is more cost-effective than C and leads to better health outcomes. The rationale for extended dominance is that a decision maker who is willing to pay the dominated CER of C can better pay the lower CER of implementing B combined with D, which leads to more health effects. The figure also illustrates how easy it can be to misrepresent the efficiency of a program by comparing it with the wrong comparator. An ICER that compares a new intervention A to an obsolete and irrelevant comparator X may make A *appear* favorable (the slope of the dotted line connecting both points is lower than the slope of the dashed cost-effectiveness threshold line through the origin), but in fact both strategies A and X are dominated. In Figure 5.2, the relevant ICERs to be reported and considered by the decision maker are B versus O and D versus B. Compared with the threshold, B versus O is clearly cost-effective, whereas D versus B is not. In general, the intervention with the smallest slope is the most cost-effective one and should be implemented first, followed by the one with the

smallest slope starting from that intervention onward, and so on.

Note: The slope of the lines in Fig. 5.2 are all CERs. Average CERs represent the cost per health effect achieved by a program (e.g., the slope of OA). Marginal CERs represent the change in cost-effectiveness when the scale of a program is varied (e.g., AA′). Incremental CERs represent the extra cost per extra health effect of one program versus another (e.g., OB or DB). An ACER is a particular case of an ICER (i.e., a comparison with the "do nothing" scenario)

5.2.4.2 Net Benefits

Some studies prefer to express cost-effectiveness results as "net benefits" rather than as ICERs because the latter is a ratio instead of a single number, which has a number of analytical disadvantages and is more difficult to interpret. For instance, as a ratio, the ICER does not give an indication of the scale of the programs being considered. Also, ICERs falling in the southeast and northwest quadrants of the CE plane (Fig. 5.1) will have the same (negative) sign, although we would want to adopt the former (more effective/less costly) intervention but not the latter (less effective/more costly). Moreover, for statistical analyses, net benefits can be easier to work with than ICERs. Net benefits incorporate the threshold willingness-to-pay value

for a gain in health outcome (to which CERs otherwise need to be compared in order to assess whether they are too expensive). Net benefit is calculated by subtracting the incremental costs of the programme (ΔC) from the monetary equivalent (WTP[E]) of the achieved incremental health gain (ΔE) it would generate. A value above zero indicates a net gain, and a negative value indicates a net loss.

$$Net\ \text{benefit} = WTP(E) * \Delta E - \Delta C$$

The net benefit approach resembles CBA, which also expresses both costs and effects in monetary values. CBA (see Chap. 4), however, typically allows the patient to do the valuing of the health effects, whereas net benefits usually represent a valuation of health gains by the general public (welfarism vs. extra-welfarism; see Chap. 9). If so, CBA implies an overall valuation of all the specific consequences of the program (including highly particular effects on individual patients' quality of life, such as improved social life; ability to work, parent, participate in sports; and the degree to which these particular aspects matter to a patient), whereas net benefits based on social valuations only provide generic values for the particular health consequence that was measured.

5.2.5   Uncertainty

A combination of inputs on costs, outcomes, and probabilities leads to a point estimate of the incremental cost per outcome of one intervention versus another (as illustrated in Fig. 5.1). But the accuracy of this estimate depends on the degree of uncertainty that is embodied in the underlying observations and calculations, and it would be misleading not to report this uncertainty in the final results.

Three sources of uncertainty can be distinguished [61]. The first is parameter uncertainty: uncertainty in the input variables that are used. This is mainly the result of sampling and measurement error, in that the observed estimates are at best only an approximation of the "real" value of a parameter. Second, there will be *structural* uncertainty, related to uncertainty in the model-ing approach or the trial design. For instance, are any disease outcomes ignored in the model or the trial? Are disease outcomes or treatment outcomes really independent, as is assumed in the analysis, or do different arms of the trial or branches of the decision tree in reality interact? And finally, there is *methodological* uncertainty. Are the methods used in the CEA sufficient to measure the costs and outcomes of an intervention? For instance, is the outcome chosen the most relevant for measuring health gain in a particular area? Is it sensitive enough to reflect meaningful changes in outcomes? Do discount rates represent social time preferences? Should indirect costs be considered and, if so, how? This more general type of uncertainty about how to measure the efficiency of an intervention cannot easily be solved and is most relevant to the correct interpretation of the results (see Sect. 5.3).

The effect of parameter and structural uncertainty can mostly be analyzed via "sensitivity analysis" (see Chap. 7), exploring the impact on the estimated CER of making different assumptions in terms of models and parameters. Structural uncertainty can be addressed by exploring the effect of different model structures. Parameter uncertainty can be dealt with by changing the value of particular inputs. In univariate, deterministic sensitivity analysis, alternative values are used for an individual key model parameter (e.g., the price of a drug). In multivariate, deterministic sensitivity analysis, the effect of changing many assumptions at the same time is explored (also called a "scenario analysis"). These alternative values are still determined by the analyst.

In probabilistic sensitivity analysis, statistical distributions are added to variables from which random values are drawn (e.g., 10,000 random picks). These iterations lead to a "cloud" of cost-effectiveness estimates (10,000 estimates) across the four quadrants of the CEA plane, which gives a general indication of the location of the "real" ICER, given the statistical distributions of the variables used. The magnitude of this cloud indicates the extent of the uncertainty that is embodied in the ICER. It also shows whether mainly outcomes or costs are uncertain, or both. Figure 5.3 illustrates
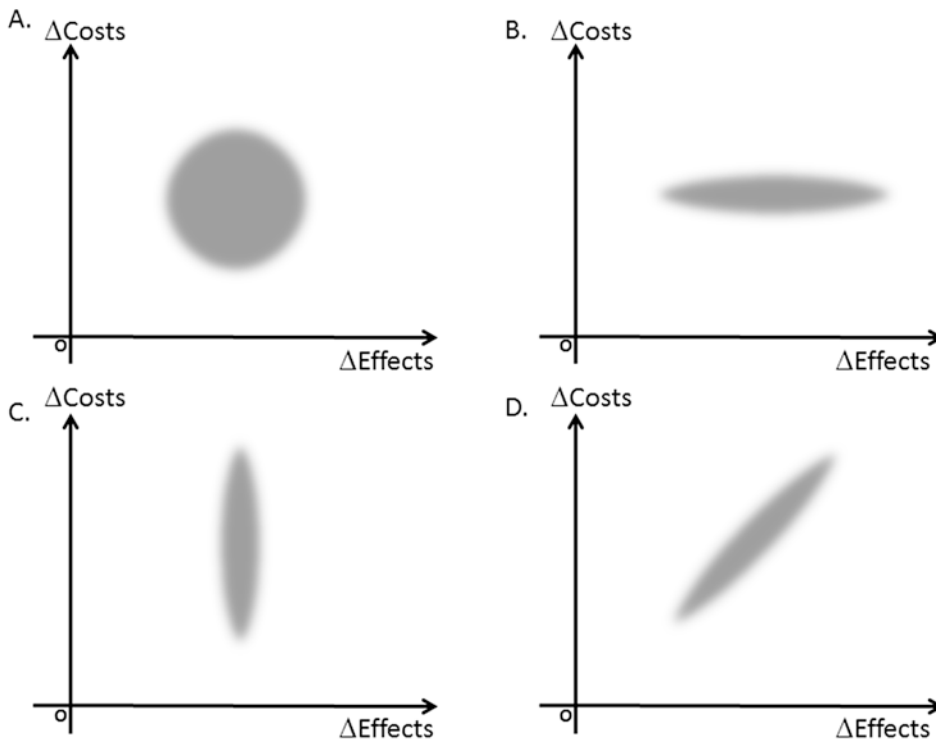
**Fig. 5.3** Cost-effectiveness clouds

this. In panel A, costs and effects are equally uncertain. A cloud resembling a horizontal ellipse (panel B) indicates that the variation in outcomes is greater than the variation in cost estimates. In panel C, effects are more certain than costs. Costs and effects can also be correlated. Panel D represents a situation where costs and effects are equally uncertain but positively correlated.

Sensitivity analysis can be used to identify the main drivers of the results and the inputs for which further research can reduce uncertainty. For instance, it can demonstrate that the most influential variable in the cost-effectiveness of an antidepressant is the effectiveness of the drug in the patient group younger than 60 years of age. Consequently, those who interpret the CEA need to judge the certainty of the particular value of that parameter that was used in the study. If there is substantial uncertainty about this estimate, an additional "value of information analysis" (VOI) can be performed to establish the monetary value of acquiring additional information (i.e., certainty) on that specific parameter, which can consequently be compared with the extra research

cost of obtaining it [62, 63] (see Chap. 7). VOI can be used to aid decision makers by demonstrating how much it would cost to reduce the uncertainty surrounding the resource allocation decision (e.g., by increasing the sample size), and whether the cost is worth incurring, versus making that decision on the basis of the presently available information [64]. For an example of the use of VOI in the mental health field, see the work by McCollister and colleagues [47].

A convenient way to graphically represent the uncertainty involved in an ICER is the "cost-effectiveness acceptability curve" (CEAC), which is illustrated in Fig. 5.4. CEACs are a different way of representing cost-effectiveness clouds and visualize, for every willingness-to-pay threshold per outcome gained, the proportion of ICER estimates that would fall below that threshold; put another way, CEACs show the probability that the net monetary benefit is greater than zero at each of a range of potential willingness-to-pay values. This point is illustrated in Fig. 5.4, where 50% of the estimated ICERs of one intervention over another lies below the threshold willingness-to-pay

Fig. 5.4 Cost-
effectiveness
acceptability curve
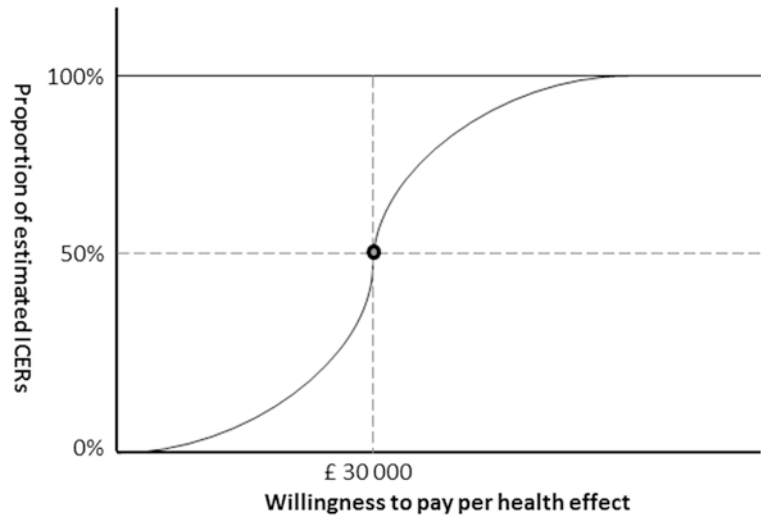
value of £30,000. This means that a decision maker has a 50% chance that the intervention will offer good value for the money and a 50% chance that it will be too expensive, relative to that specific threshold value.

To conclude this section, most countries have developed practical guidelines for analysts on how to handle the technical assumptions and controversies related to quantifying costs, effects, uncertainty, and results [65, 66].

5.3   Interpretation

Once presented with the results of a CEA, a decision maker is faced with the task of assessing and interpreting the evidence at hand. The decision maker must assess the quality and also the usefulness of the evidence. We address these issues in turn.

First, what is the quality of the study in assessing the real "value for money" of the intervention? Is uncertainty properly accounted for? Are the options under evaluation clearly defined and described? Are differences in reported costs and effects between interventions fully attributable to the interventions or also to unreliable or invalid methodologies, which is less desirable? Are important categories of costs neglected? As

mentioned, economic evaluations of mental health interventions may be particularly sensitive to the perspective adopted in the analysis, and an atypical cost profile often occurs in mental health. Such broader costs can be estimated, but often with a degree of uncertainty. Are all relevant outcomes captured? CEA uses specific effect measures that may focus on only one aspect of an illness and neglect other important outcomes. It may also fail to capture the adverse effects of a treatment.

Second, assessing whether an intervention is cost-effective (i.e., it is "worth it") requires a benchmark – a cost-per-effect threshold – that distinguishes health benefits that come at a "reasonable" cost from those that are excessively costly. Benchmarks or threshold values for a life year in full health (a QALY) exist in several countries, but typically not for condition-specific health outcomes. This immediately brings us back to the main weakness of CEA in providing information on efficient resource allocation. CEA allows assessments of efficiency at a local level, within the budget available for a condition or to achieve a particular outcome. But ultimately, a more general idea of the value of one particular type of effect (e.g., one depressive episode) in the overall picture of health and well-being is still required to assess whether costs are acceptable.

How many other health gains, products, or services is a society willing to give up for a gain in one particular mental health outcome? This limitation of CEA is relevant in the context of mental health, where there remains a major challenge to obtain funding that is proportionate to the disease burden associated with mental health disorders. Mental health interventions are often seen by policymakers as less important than physical health interventions, as the prevailing conception of health and sickness is still predominantly a biomedical one. CEA, constrained to particular mental health specialties, cannot address issues of allocative efficiency across the wider spectrum of healthcare specialties.

Last but surely not least, an efficient allocation of the available resources will maximize achievable health effects under budget constraints. But this outcome is not necessarily the most desirable one from a social or an ethical perspective. It does not acknowledge the relation of health programs with other important objectives of healthcare, including tackling health inequities; promoting respect for individual autonomy, dignity, and patient preferences; personal responsibility; solidarity with the worst-off groups in society; or even bioethical considerations about the moral desirability of particular technologies [67]. There may be good reasons why a less efficient program still deserves funding, or why an efficient strategy is not desirable. However, CEA would indicate that accommodating and upholding other ethical values would come at a higher opportunity cost. This point is discussed in further detail in Chap. 10.

5.4   Conclusion

CEA is of most use in situations where (a) a budget holder needs to make allocation decisions among a number of options within a particular clinical field (or has "ring-fenced" money to spend), and (b) there is a clear measure of success. It is increasingly used to complement evidence of the efficacy and effectiveness of interventions in order to demonstrate that the

costs of an intervention are also proportionate to the gains achieved. In the context of mostly fixed and pressurized healthcare budgets, these considerations of efficiency become increasingly relevant. Given its increasing effect on decision-making, it is important that individuals who work in the field of mental health policy are familiar with the primary components and assumptions of CEA, the complexities inherent to the methodology, and the particular challenges that occur when it is applied to the context of mental health.

> **Key Messages**
>
> - CEA compares the costs of implementing a mental health program with its achieved outcome. In contrast to CBA or CUA, this outcome is defined in terms of natural units that are specifically relevant to a particular disease area.
> - CERs provide an indication of the efficiency of resource allocation within a particular disease area. What do competing programs cost per health effect achieved? Or, vice versa, per amount invested in a program, how much improvement in health effects can be "bought"?
> - Results are sensitive to the costing perspective that is adopted and to whether all relevant costs are considered. As atypical cost patterns may emerge in mental healthcare, this is an important point to highlight.
> - Cost-effectiveness estimates embody large uncertainties, but methods exists to account for these. The quality of a study can often be judged by the extent to which this uncertainty is dealt with.
> - Cost-effectiveness estimates provide useful but nonetheless complex information to an already difficult decision-making process, and they do not "make decisions." To avoid oversimplification, attention must be paid to the correct normative interpretation of study results.

References

1. Gold M, Siegel JE, Russell LB, Weinstein MC. Cost-effectiveness in health and medicine. New York: Oxford University Press; 1996.
2. Drummond M, Sculpher M, Torrance GW, O'Brien B, Stoddard G. Methods for the economic evaluation of health care programmes. Oxford: Oxford University Press; 2005.
3. Edejer T, Baltussen R, Adam T, Hutubessy R, Acharya A, Evans D, et al. WHO guide to cost-effectiveness analysis. Geneva: World Health Organization; 2003.
4. Clement FM, Harris A, Li JJ, Yong K, Lee KM, Manns BJ. Using effectiveness and cost-effectiveness to make drug coverage decisions: a comparison of Britain, Australia, and Canada. JAMA. 2009;302(13):1437–43.
5. Dakin H, Devlin N, Feng Y, Rice N, O'Neill P, Parkin D. The influence of cost-effectiveness and other factors on nice decisions. Health Econ. 2014; 24(10):1256–71.
6. Luyten J, Naci H, Knapp M. Economic evaluation of mental health interventions: an introduction to cost-utility analysis. Evidence-based mental health. 2016;19(2).
7. Shearer J, McCrone P, Romeo R. Economic evaluation of mental health interventions: a guide to costing approaches. PharmacoEconomics. 2016;34(7):651–64.
8. Thomas CM, Morris S. Cost of depression among adults in England in 2000. Br J Psychiatry J Ment Sci. 2003;183:514–9.
9. Luyten J, Beutels P. Costing infectious disease outbreaks for economic evaluation: a review for hepatitis A. PharmacoEconomics. 2009;27(5):379–89.
10. Mulhern B, Mukuria C, Barkham M, Knapp M, Byford S, Soeteman D, et al. Using generic preference-based measures in mental health: psychometric validity of the EQ-5D and SF-6D. Br J Psychiatry. 2014;205(3):236–43.
11. Brazier J. Is the EQ–5D fit for purpose in mental health? Br J Psychiatry. 2010;197(5):348–9.
12. Saarni SI, Viertiö S, Perälä J, Koskinen S, Lönnqvist J, Suvisaari J. Quality of life of people with schizophrenia, bipolar disorder and other psychotic disorders. Br J Psychiatry. 2010;197(5):386–94.
13. Papaioannou D, Brazier J, Parry G. How valid and responsive are generic health status measures, such as EQ-5D and SF-36, in schizophrenia? A systematic review. Value Health. 2011;14(6):907–20.
14. Hastrup LH, Nordentoft M, Hjortheøj C, Gyrd-Hansen D. Does the EQ-5D measure quality of life in schizophrenia? J Ment Health Policy Econ. 2011;14(4):187–96.
15. Priebe S, Savill M, Wykes T, Bentall RP, Reininghaus U, Lauber C, et al. Effectiveness of group body psychotherapy for negative symptoms of schizophrenia: multicentre randomised controlled trial. Br J Psychiatry. 2016;209(1):54–61.
16. Hollinghurst S, Peters TJ, Kaur S, Wiles N, Lewis G, Kessler D. Cost-effectiveness of therapist-delivered online cognitive-behavioural therapy for depression: randomised controlled trial. Br J Psychiatry. 2010;197(4):297–304.
17. Guy W. Clinical Global Impressions Scale (CGI). In: Rush AJ, editor. Handbook of psychiatric measures. Washington, DC: American Psychiatric Association; 2000.
18. King D, Knapp M, Thomas P, Razzouk D, Loze JY, Kan HJ, et al. Cost-effectiveness analysis of aripiprazole vs standard-of-care in the management of community-treated patients with schizophrenia: STAR study. Curr Med Res Opin. 2011;27(2):365–74. 10p.
19. Kay SR, Fiszbein A, Opler LA. The Positive and Negative Syndrome Scale (PANSS) for schizophrenia. Schizophr Bull. 1987;13(2):261–76.
20. Priebe S, Savill M, Wykes T, Bentall R, Lauber C, Reininghaus U, et al. Clinical effectiveness and cost-effectiveness of body psychotherapy in the treatment of negative symptoms of schizophrenia: a multicentre randomised controlled trial. Health Technol Assess. 2016;20(11):1–100.
21. Tandon R, Devellis RF, Han J, Li H, Frangou S, Dursun S, et al. Validation of the Investigator's Assessment Questionnaire, a new clinical tool for relative assessment of response to antipsychotics in patients with schizophrenia and schizoaffective disorder. Psychiatry Res. 2005;136(2–3):211–21.
22. American Psychiatric Association. Diagnostic and statistical manual of mental disorders. 4th ed. Washington, DC: American Psychiatric Association; 1994.
23. Hastrup LH, Kronborg C, Bertelsen M, Jeppesen P, Jorgensen P, Petersen L, et al. Cost-effectiveness of early intervention in first-episode psychosis: economic evaluation of a randomised controlled trial (the OPUS study). Br J Psychiatry. 2013;202(1):35–41.
24. Beck A, Steer R, Brown G. Manual for the beck depression inventory: the psychological corporation. 1987;19(2).
25. Beck AT, Erbaugh J, Ward CH, Mock J, Mendelsohn M. An inventory for measuring depression. Arch Gen Psychiatry. 1961;4(6):561.
26. Kuyken W, Byford S, Taylor RS, Watkins E, Holden E, White K, et al. Mindfulness-based cognitive therapy to prevent relapse in recurrent depression. J Consult Clin Psychol. 2008;76(6):966–78.
27. Maljanen T, Knekt P, Lindfors O, Virtala E, Tillman P, Harkanen T, et al. The cost-effectiveness of short-term and long-term psychotherapy in the treatment of depressive and anxiety disorders during a 5-year follow-up. J Affect Disord. 2016;190:254–63.
28. Zigmond AS, Snaith RP. The hospital anxiety and depression scale. Acta Psychiatr Scand. 1983;67(6):361–70.
29. Romeo R, Knapp M, Banerjee S, Morris J, Baldwin R, Tarrier N, et al. Treatment and prevention of depression after surgery for hip fracture in older people: cost-effectiveness analysis. J Affect Disord. 2011;128(3):211–9.

30. Alexopoulos GS, Abrams RC, Young RC, Shamoian CA. Cornell scale for depression in dementia. Biol Psychiatry. 1988;23(3):271–84.

31. Banerjee S, Hellier J, Romeo R, Dewey M, Knapp M, Ballard C, et al. Study of the use of antidepressants for depression in dementia: the HTA-SADD trial – a multicentre, randomised, double-blind, placebo-controlled trial of the clinical effectiveness and cost-effectiveness of sertraline and mirtazapine. Health Technol Asses. 2013;17(32):1–166.

32. Goldberg DP, Hillier VF. A scaled version of the General Health Questionnaire. Psychol Med. 1979;9(1):139–45.

33. Woods R, Bruce E, Edwards R, Elvish R, Hoare Z, Hounsome B, et al. REMCARE: reminiscence groups for people with dementia and their family caregivers – effectiveness and cost-effectiveness pragmatic multicentre randomised trial. Health Technol Asses. 2012;16(50):1–121.

34. First MB, Spitzer RL, Gibbon M, Williams JBW. The structured clinical interview for DSM-IV Axis I disorder with psychotic screen. New York: New York Psychiatric Institute; 1995.

35. Kuyken W, Hayes R, Barrett B, Byng R, Dalgleish T, Kessler D, et al. Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (PREVENT): a randomised controlled trial. Lancet. 2015;386(9988):63–73.

36. Cummings JL, Mega M, Gray K, Rosenberg-Thompson S, Carusi DA, Gornbein J. The neuropsychiatric inventory: comprehensive assessment of psychopathology in dementia. Neurology. 1994;44(12):2308.

37. D'Amico F, Rehill A, Knapp M, Lowery D, Cerga-Pashoja A, Griffin M, et al. Cost-effectiveness of exercise as a therapy for behavioural and psychological symptoms of dementia within the EVIDEM-E randomised controlled trial. Int J Geriatr Psychiatry. 2016;31(6):656–65. 10p.

38. Cohen-Mansfield J. Measurement of inappropriate behavior associated with dementia. J Gerontol Nurs. 1999;25(2):42–51.

39. Chenoweth L, King MT, Jeon YH, Brodaty H, Stein-Parbury J, Norman R, et al. Caring for Aged Dementia Care Resident Study (CADRES) of person-centred care, dementia-care mapping, and usual care in dementia: a cluster-randomised trial. Lancet Neurol. 2009;8(4):317–25.

40. Logsdon RG, Gibbons LE, McCurry SM, Teri L. Assessing quality of life in older adults with cognitive impairment. Psychosom Med. 2002;64(3):510–9.

41. D'Amico F, Rehill A, Knapp M, Aguirre E, Donovan H, Hoare Z, et al. Maintenance cognitive stimulation therapy: an economic evaluation within a randomized controlled trial. J Am Med Dir Assoc. 2015;16(1):63–70.

42. Orgeta V, Leung P, Yates L, Kang S, Hoare Z, Henderson C, et al. Individual cognitive stimulation therapy for dementia: a clinical effectiveness and cost-effectiveness pragmatic, multicentre, randomised controlled trial. Health Technol Asses. 2015;19(42):1–108.

43. Rosen WG, Mohs RC, Davis KL. A new rating-scale for Alzheimers-disease. Am J Psychiatry. 1984;141(11):1356–64.

44. Patel A, Knapp M, Romeo R, Reeder C, Matthiasson P, Everitt B, et al. Cognitive remediation therapy in schizophrenia: cost-effectiveness analysis. Schizophr Res. 2010;120(1–3):217–24. 8p.

45. McLellan AT, Luborsky L, Woody GE, O'Brien CP. An improved diagnostic evaluation instrument for substance abuse patients. The Addiction Severity Index. J Nerv Ment Dis. 1980;168(1):26–33.

46. Dennis ML, Titus JC, White M, Unsicker J, Hodgkins D. Global Appraisal of Individual Needs (GAIN): administration guide for the GAIN and related measures. Bloomington: Chestnut Health Systems; 2003.

47. McCollister KE, French MT, Freitas DM, Dennis ML, Scott CK, Funk RR. Cost-effectiveness analysis of Recovery Management Checkups (RMC) for adults with chronic substance use disorders: evidence from a 4-year randomized trial. Addiction. 2013;108(12):2166–74.

48. Olmstead TA, Sindelar JL, Easton CJ, Carroll KM. The cost-effectiveness of four treatments for marijuana dependence. Addiction. 2007;102(9):1443–53.

49. Beck AT, Steer RA. Manual for the Beck scale for suicide ideation. San Antonio: The Psychological Corporation; 1991.

50. van Spijker BA, Majo MC, Smit F, van Straten A, Kerkhof AJ. Reducing suicidal ideation: cost-effectiveness analysis of a randomized controlled trial of unguided web-based self-help. J Med Internet Res. 2012;14(5):e141.

51. Petrou S, Gray A. Economic evaluation alongside randomised controlled trials: design, conduct, analysis, and reporting. BMJ. 2011;342:d1548.

52. Zwarenstein M, Treweek S, Gagnier JJ, Altman DG, Tunis S, Haynes B, et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ. 2008;337:a2390.

53. Jones PB, Barnes TR, Davies L, Dunn G, Lloyd H, Hayhurst KP, et al. Randomized controlled trial of the effect on quality of life of second- vs first-generation antipsychotic drugs in schizophrenia: Cost Utility of the Latest Antipsychotic Drugs in Schizophrenia Study (CUtLASS 1). Arch Gen Psychiatry. 2006;63(10):1079–87.

54. Stroup TS, McEvoy JP, Swartz MS, Byerly MJ, Glick ID, Canive JM, et al. The National Institute of Mental Health Clinical Antipsychotic Trials of Intervention Effectiveness (CATIE) project: schizophrenia trial design and protocol development. Schizophr Bull. 2003;29(1):15–31.

55. Lewis S, Lieberman J. CATIE and CUtLASS: can we handle the truth? Br J Psychiatry. 2008;192(3):161–3.

56. Gillies C, Freemantle N, Grieve R, Sekhon J, Forder J. Advancing quantitative methods for the evaluation of complex interventions. In Raine, R., Fitzpatrick, R.,

Barratt, H., Bevan, G., Black, N., Boaden, R. et al. Challenges, solutions and future directions in the evaluation of service innovations in health care and public health. Health Serv Deliv Res. 2016;4(16):37–54.

57. Briggs A, Claxton C, Sculpher M. Decision modelling for health economic evaluation. Oxford: Oxford University Press; 2006.

58. Detsky AS, Naglie G, Krahn MD, Naimark D, Redelmeier DA. Primer on medical decision analysis: part 1 – getting started. Med Decis Making Int J Soc Med Decis Making. 1997;17(2):123–5.

59. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Making. 1993;13:322–38.

60. Oliver A. A normative perspective on discounting health outcomes. J Health Serv Res Policy. 2013;18(3):186–9.

61. Bilcke J, Beutels P, Brisson M, Jit M. Accounting for methodological, structural, and parameter uncertainty in decision-analytic models: a practical guide. Med Decis Making Int J Soc Med Decis Making. 2011;31(4):675–92.

62. Drummond M, Sculpher MJ, Torrance GW, O'Brien B, Stoddart GL. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford: Oxford University Press; 2005.

63. Ades AE, Lu G, Claxton K. Expected value of sample information calculations in medical decision modeling. Med Decis Making. 2004;24(2):207–27.

64. Wilson EC. A practical guide to value of information analysis. PharmacoEconomics. 2015;33(2):105–21.

65. KCE. Guidelines for pharmacoeconomic evaluations in Belgium In: Centre BHCK, editor. KCE reports 78C. Brussels. 2008.

66. National Institute for Health and Care Excellence. Guide to the methods of technology appraisal. London: National Institute for Health and Clinical Excellence; 2013. June. 2008. Report No.

67. Schokkaert E. How to introduce more (or better) ethical arguments in Hta? Int J Technol Assess Health Care. 2015;31(3):111–2.