Rationalization of innovation: the role of health-economic evaluation in improving the efficiency of mental health care

Joran Lokkerbol

Rationalization of innovation:
the role of health-economic evaluation in improving the efficiency of mental health care

Joran Lokkerbol

VRIJE UNIVERSITEIT

Rationalization of Innovation: The role of health-economic evaluation in improving the efficiency of mental health care

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. F.A. van der Duyn Schouten,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de Faculteit der Psychologie en Pedagogiek
op woensdag 6 mei 2015 om 15.45 uur
in de aula van de universiteit,
De Boelelaan 1105

door

Joran Lokkerbol

geboren te Zijpe

promotoren: prof.dr. H.F.E. Smit
                      prof.dr. W.J.M.J. Cuijpers

copromotoren: dr. R.M. Weehuizen
                      dr. C. Mihalopoulos

Contents

General Introduction

General Introduction

This thesis considers the rationalization of innovation in mental health care by looking at available and new health economic methodologies that help to improve the efficiency of health care systems. The thesis is divided in three parts:

> 1. The need for a cost-effective mental health care system
> 2. Established methods for assessing the cost-effectiveness of interventions
> 3. New approaches to improving the cost-effectiveness of health care systems.

Pursuing the goal of rationalization of innovation should begin with outlining the need for innovation. The high disease burden of the mental disorders (via losses in quality of life and the indirect association with premature death) and the large impact on society in terms of health care costs and productivity losses make it a worthwhile goal to pursue the alleviation of disease burden (Hoffman et al 2008; Rehm et al 2009; Cuijpers et al 2012). However, currently available mental health services only manage to alleviate a relatively small fraction of the disease burden (Andrews et al 2004; Chisholm et al 2004a), partly stemming from the inefficiency of current health care (Andrews et al 2004; Vos et al 2010). Furthermore, the 'gap' between what treatment can offer and the needs of many is likely to widen (World Health Organization, 2008). Worldwide, populations are aging, causing a relative increase in the (older) population. In developed regions, the relative increase in the older population is accompanied by a simultaneous relative decrease in the (working) population that ought to provide the health care (United Nations, 2013). As health care has to be provided and funded by a relatively smaller proportion of the population, this is likely to impact on already stretched health care systems, both in mental and somatic health care.

  Another reason to innovate mental health care, is the prominent position that mental disorders have on (non-fatal) burden of disease rankings (Vos et al 2012). Furthermore, as infectious diseases are being managed better, their relative contribution to the global disease burden diminishes. Non-communicable diseases such as mental disorders therefore have an increasingly large share in the global disease burden, particularly in high-income countries.

Overall, the economic downturn compounds and exacerbates these demographic and epidemiological transitions, by putting pressure on resources needed to manage the high disease burden associated with mental disorders.

To better equip our health care system to lessen the disease burden stemming from mental disorders, it is of vital importance that we start developing new methodologies and strategies that enhance the cost-effectiveness of our health care systems while simultaneously taking into account the limited resources that we have, both financially and in terms of human resource capacity.

From a policy-making and health economic perspective, limited therapists' time and tight financial budgets should be invested in areas where these have the greatest leverage on population health. Within the context of mental health care, this means that efforts should be directed at populations where the disease burden and economic costs can be dealt with most (cost-)effectively, in ways that respect resource constraints. Furthermore, with the ongoing development of new interventions and health technologies such as E-health (Riper et al 2010) or M-health (mobile health) (Harrison et al 2011), as well as evolving preferences of patient groups and health care providers, health economic evaluation can continuously contribute by synthesizing available evidence and incorporating this in the decision making process.

When one is convinced of the importance of on-going innovation, the next step is to use the tools at hand to guide the process of health care innovation. A range of economic evaluation methods has been used to support decision making in health care, often by comparing a new treatment with care as usual in terms of its relative improvements of cost-effectiveness, and recommending adding the new intervention to the existing package of interventions when the new intervention is considered to be cost-effective (for example, see Holman et al 2011; Annemans et al 2014; Aragonèsa et al 2014; Wiles et al 2014). However, this process of adding new interventions generally leads to increasing health care expenditure overall, because offering more interventions (even if they are cost-effective) in general requires a larger health care budget (for example, see Sindern et al 2010). Under the current economic downturn where budgets are not easily increased, this route to innovation seems to lose its practical value. Alternative methods, explicitly taking budget constraints into account, could help to fill this gap. The last part of this thesis presents some new health economic approaches that meet these requirements.

Before presenting the outline of this thesis, we will now first introduce some terminology to clarify the key concepts and principles underlying this thesis.

Economics

As much of the work in this thesis is based on economic principles, it is worthwhile to give a brief overview of the main principles underlying this discipline.

This thesis is primarily concerned with increasing the allocative efficiency of scarce resources within the health care system (i.e. which services/goods should be provided with societal scarce resources (Mihalopoulos 2011)). The economic principles guiding this process of increasing the allocative efficiency, are 'opportunity costs', 'marginal analysis' and 'benefit'.

When trying to determine how scarce resources are used best, alternative ways in which these available resources can be used have to be compared. Resources can only be used once, so choosing to allocate budget to the treatment of depression means that this budget can no longer be used for the treatment of anxiety disorder, alcohol disorder, schizophrenia, etc. The 'opportunity cost' of allocating a budget to the treatment of depression refers to the value foregone that could have been obtained when using this budget alternatively (e.g. for treatment of alcohol disorder). The principle of 'opportunity costs' helps us in determining *where* to invest, by using our budget in areas where the 'opportunity cost' is minimal.

Next to determine *where* to spend budget, it is important to know *how much* budget to spend. This is where the principle of 'marginal analysis' comes in. Based on the idea that the additional value of spending additional budget decreases (first treating patients who benefit most from a treatment, then treating patients who benefit slightly less from this treatment, etc.), the principle of 'marginal analysis', in combination with the principle of 'opportunity costs', imply that budget should be increased up to the point where the additional benefit of increasing the budget is equal to the 'opportunity cost' of this budget. From that point on, the 'opportunity cost', or potential value that can be obtained when spending budget on an alternative, is higher than the additional value achieved when spending more budget on the same indication.

The last thing needed to determine a theoretically efficient allocation, is the concept of 'benefit'. As 'opportunity costs' and 'marginal analysis' are concerned with *where* to spend *how much* budget, 'benefit' defines the outcome we are trying to optimize. After all, determining *where* and *how much* budget to spend within (mental) health care depends on the outcome of interest (e.g. total health in the

population, the monetary value associated with health outcomes, some measure of equity, etc.).

Economic evaluation techniques
Depending on the outcome of interest, different economic evaluation techniques can be applied to evaluate alternative uses of resources in their ability to improve allocative efficiency. The three most commonly used techniques are cost-benefit analysis (CBA), cost-utility analysis (CUA) and cost-effectiveness analysis (CEA).

CBA uses a monetary outcome measure (benefit), which requires outcomes (such as health, education or voluntary work) to be valued in monetary terms. For some benefits (e.g. fewer days out of work) it is relatively easy to compute the monetary value. The (euro or dollar) value of other benefits is less easy to compute and may rely on some willingness-to-pay (WTP) technique to establish their value. Although determining the WTP is not straightforward, the advantage of CBA is the ability to combine diverse outcome measures, impacting on diverse sectors, into one decision-making framework (Mihalopoulos 2011).

CUA uses a generic health-related quality of life indicator as an outcome measure, quality-adjusted life years (QALYs), in evaluating different alternatives. The advantage is that using the QALY makes CUA suitable for comparing alternatives across different health conditions.

CEA is based on a disease-specific 'effect' measure (e.g. change in depressive symptoms as measured by the Beck Depression Inventory), thereby limiting its use to comparing alternatives within similar diagnostic groups, but increasing the usability of such comparisons within a diagnostic target group.

In health care markets, the market cannot be relied upon to come to an optimal resource allocation due to fundamental market failures (e.g. asymmetry in information, or high barriers of entry (Hsiao and Heller 2007)). Therefore, schools of thought arose within economics to address the problem of efficient allocation. Different economic evaluation techniques relate to different schools of thought within economics. It is beyond the scope of this thesis to provide a detailed description of the different schools of thought, but interested readers could consult the literature on 'welfarism' (Sugden and Williams 1978; Drummond 1981; Mihalopoulos 2011), strongly linked to CBA, 'extra-welfarism' (Culyer 1990; Hurley 1998, 2000; Brouwer and Koopmanschap 2000; Birch and Donaldson 2003; Mihalopoulos 2011), related to CUA and CEA, and the 'decision-making school' (Sugden and Williams 1978; Carter 2001; Carter et al 2008; Mihalopoulos 2011), argued to be consistent with CBA, CUA and CEA.

The glossary at the end of this dissertation describes the relevant (health) economic terms used throughout this thesis.

Innovation

In the context of this thesis, health care innovation refers to improving health care systems. Innovation is the response to so-called 'drivers of innovation', such as the introduction of new technologies, the changing preferences and demands of health care users, or changing financial incentives or constraints. In this context, health care innovation is seen as the ongoing process of change in the organisation of health care systems and their actors. These actors (health care users, health care providers, health care authorities and health care financiers) are part of the health care system and the innovation process, and their perspectives are therefore explicitly taken into account (we return to this shortly).

To be specific, we define a health care system as a set of interventions aimed at alleviating the disease burden of its target groups. These interventions are further characterized by:
- the number of patients treated (coverage),
- patients' compliance (adherence),
- per patient health improvement (effectiveness), and
- per patient intervention costs (resource use).

Each intervention, which could range from self-management, e-health, psychological care, pharmacotherapy or combinations thereof, contributes to and affects the health care system, which, when judged by its ability to alleviate the disease burden of its target group, needs to be:
- appropriate in the eyes of the health care providers
- acceptable from the perspective of the health care users
- effective in terms of clinical effectiveness, and
- affordable to better guarantee its economic sustainability.

In this thesis, health care innovation refers to the improvement of health care systems across these dimensions (appropriateness, acceptability, effectiveness, affordability) by changing the health care system's underlying package of interventions, for example by adding new interventions or rearranging the mix of existing interventions.

At this point, it is important to note that in this context an improvement both has an absolute and a relative meaning. Improvement is absolute in the sense that we always prefer a health care system that is more appropriate, more acceptable, more effective or more affordable. Improvement is relative in the sense that these qualifiers are interrelated and an improvement in one area could lead to a worsening in another area. For example, changing an intervention in order to improve health care users' compliance could increase a health care systems' acceptability, while making it less affordable. In that case, the health care system has improved from the patients' perspective, but has become worse (more expensive) from the financier's perspective. It is not immediately clear which scenario is preferred from an overall policy perspective, where such decisions need to be made. These interdependencies between the various perspectives make improving the health care system a non-straightforward process. First of all, health care innovation requires taking every perspective into account: the patient-perspective (acceptability), the health care provider's perspective (appropriateness) and the financier's perspective (affordability and effectiveness; combined in cost-effectiveness), which are all relevant for the overarching policymaking perspective, which could entail national policymakers, but also policymakers on regional, local, or institutional level. In line with this, health care innovation requires addressing each of the four criteria (appropriateness, acceptability, effectiveness, affordability) in order for a health care system to be successful in alleviating disease burden and improving health. That is, a health care system that is appropriate, acceptable and evidence-based but not affordable will not be successful in reducing the disease burden in an optimal way. In general, whenever one or more of these four corner stones are not taken care of, the health care system is unlikely to be sustainable over the longer term. Innovation, if it pursues durable improvement of the health gains, should take into account every perspective, as change will only be supported when taking the interests of every stakeholder into account.

In this thesis, health care innovation refers to improving the extent in which a health care system is successful in improving population health given resource constraints. Innovation can involve adding more cost-effective interventions, but also creating a different mix of interventions, which ultimately leads to improvement of coverage, adherence, effectiveness and economic affordability.

Rationalization

Where *innovation* refers to improving population health in the context of monetary constraints, *rationalization* refers, in the context of this thesis, to the structured approach of combining the best available evidence from the fields of epidemiology, economics and clinical psychology and pharmacology (in line with Vos et al 2010) to empirically and theoretically support these innovations.

In chapters 3–6 we combine the evidence from these different fields into health economic simulation models, thereby quantifying the overall costs and effects of health care, but also offering the possibility to quantify the impact on cost-effectiveness when changing the health care system, for example by adding new interventions, or changing the current intervention mix. Health economic models, as opposed to trial-based studies (e.g. chapter 2), allow for a longer time horizon, the possibility to compare all relevant options (rather than only the options included in a trial), and to synthesize evidence (Briggs et al 2006). Trial-based studies on the other hand, provide empirical (data-supported) estimates of the treatment effect and the associated costs. Trial based studies often provide input for health economic models. In turn, health economic models can provide input on where more information is needed, thereby providing input for research agendas on where trial-based studies are needed (Briggs et al 2006).

The use of a health economic model makes it possible for each perspective (patient, health care provider, health care financier, policymaker) to be specifically taken into account by quantifying the change in coverage, adherence, effectiveness and per-patient intervention cost as a result of innovation and translating this into population health (e.g. QALYs gained; DALYs averted) and required budget (in euro). Specific input data are required to populate the health economic simulation models. Ideally, the input into the health economic model is based on empirical trial data. However, for some input parameters, such as compliance rates for interventions, it might be hard to obtain empirical data. In such cases, estimates could be obtained using focus groups of patients or expert opinion. If this is not feasible or available, ad hoc estimates could be used of which the impact could be tested afterwards using extensive sensitivity analyses. No matter how strong the underlying evidence base, input parameters will always be subject to some degree of uncertainty. This is specifically taken into account by calculating population health and budgets not just once, but for example a 1,000 times, each time using input parameter values that are randomly drawn from a distribution of the input parameter within given upper and lower limits.

A health economic model could for example be used to compare current health care with the same current health care that is augmented with an additional intervention. But once current health care is defined, it is possible to perform all sorts of 'what-if' analyses, such as: what is the impact on cost-effectiveness when (partly) substituting face-to-face care by e-health? Or, what is the impact on cost-effectiveness when offering more prevention? Or, is it better to focus on improving coverage, or is system innovation better served when improving adherence to existing treatments?

Outline of the thesis

The overall outline of this thesis can be summarized as follows. The first part of this thesis (chapter 1) addresses the need for a well-functioning mental health care system by quantifying the non-fatal disease burden (years lived with disability, YLD) due to mental disorders in the Netherlands. Generally, only a small fraction of the overall disease burden is alleviated by the health care system (Andrews et al 2004; Chisholm et al 2004a), highlighting the importance of innovation in mental health care.

An established way of rationalizing this process of health care innovation is via economic evaluation of new interventions, either by trial-based cost-effectiveness analysis, or by combining available epidemiological, clinical and economic evidence into health economic simulation models. The second part of this thesis (chapters 2-4) presents three examples of these economic evaluations with increasing levels of sophistication. Chapter 2 presents a cost-effectiveness analysis of a four-arm randomized trial of a web-based intervention treating patients with depression and/or anxiety. Chapter 3 presents a health economic model with a one-year time horizon assessing the impact on overall cost-effectiveness of the health care system regarding alcohol disorders when offering additional online (e-health) interventions. Chapter 4 presents a multi-year health economic (Markov) model assessing the impact on overall cost-effectiveness and budget impact when using telemedicine to prevent first onset and recurrent episodes of depression.

The third part of this thesis (chapters 5-6) investigates new approaches to improving the cost-effectiveness of health care, rather than merely adding new, presumably more cost-effective, interventions. Murray et al. (2000) argue that *"addressing current allocative inefficiencies in many countries may yield substantial health gains, possibly more than identifying new technologies that will make small improvements in health"*. Following this suggestion, chapter 5 takes primary mental health care in the Netherlands as a starting point to develop an algorithm that offers strategic directions for health system improvement, specifically by changing coverage, adherence, clinical effectiveness and intervention costs of the current

intervention mix to see how the health care system can be optimized overall. Chapter 6 describes a substitution algorithm presenting intervention pairs within the current intervention mix that have potential to jointly create a health care system that generates at least as much health for equal or less budget by simultaneous investment and disinvestment. This substitution algorithm is applicable to any health care system, and is illustrated by using the health care system for major depression in the Netherlands.

Scope of the thesis

Two main restrictions apply to the scope of this thesis. First of all, health care improvement is viewed from the perspective of cost-effectiveness. Besides cost-effectiveness, policymakers need to take into account many other aspects when looking at potential improvements of the health care system, such as equity, medical ethics, feasibility, acceptability, appropriateness and strength of evidence (Vos et al 2000; Mihalopoulos et al 2011b). In this respect, this thesis is an unfinished project, starting with the health economics, which then should be complemented with a 'second-stage normative filtering' process to also take into account the various other aspects such as the impact of a health care reform on the equitable access to the health care system for all population segments.

A second restriction in the scope of this thesis relates to the costs of implementing changes in the health care system. Throughout this thesis, costs and effects of improvement options are considered, assuming that both the 'old' and the 'new' health care systems are in their 'steady state' (i.e. that the improvements have been fully implemented). This facilitates passing judgments on whether the alternative health care system is more desirable from a cost-effectiveness perspective than the current one. However, estimating the costs of changing the current health care system to its alternative falls outside the scope of this thesis.

This thesis builds on, and would not have been possible without, the extensive research performed by others. First of all, guidelines such as CONSORT (Begg et al 1996; Moher et al 2001; Schulz et al 2010) and CHEERS (Husereau et al 2013) offer invaluable contributions to the quality of the input regarding health economic simulation models. Also, extensive work on good research practices for cost-effectiveness analysis (Ramsey et al 2005) and modelling (Caro et al 2012; Roberts et al 2012; Siebert et al 2012; Karnon et al 2012; Pitman et al 2012; Briggs et al 2012; Eddy et al 2012), as well as the early and often cited contributions to the methodology of health economic evaluation (Drummond et al 2005; Briggs et al 2006), and leading examples such as the Assessing Cost Effectiveness in Prevention

(ACE-Prevention) project in Australia (Vos et al 2010) and the research related to the World Health Organization's CHOosing Interventions that are Cost-Effective (CHOICE) work programme (Tan-Torres Edejer et al 2003; Lauer et al 2003; Chisholm 2005), have been important in setting a standard in this research field. Lastly, the work of numerous individuals, such as Theo Vos (Vos et al 2004; 2005; 2010; 2012), Cathrine Mihalopoulos (Mihalopoulos 2011; Mihalopoulos et al 2011a; 2011b; 2012; 2013a; 2013b), Dan Chisholm (Chisholm et al 2004a; 2004b; 2008; 2010; Chisholm 2005; 2007), Christopher Murray (Murray and Lopez 1996; 1997; Murray et al 2000; 2001; 2004; 2012; 2013a; 2013b), has inspired the research presented in this thesis. As always, in science we stand on the shoulders of giants and fully acknowledge our indebtedness.

Part I

The need for a cost-effective mental health care system

Chapter 1. Non-fatal burden of disease due to mental disorders in the Netherlands

Chapter 1

Non-fatal burden of disease due to mental disorders in the Netherlands

Abstract

Purpose: To estimate the disease burden due to 15 mental disorders at both individual and population level.

Methods: Using a population-based survey (NEMESIS, N = 7,056) the number of years lived with disability per one million population were assessed. This was done with and without adjustment for comorbidity.

Results: At individual level, major depression, dysthymia, bipolar disorder, panic disorder, social phobia, eating disorder and schizophrenia are the disorders most markedly associated with health-related quality of life decrement. However, at population level, the number of affected people and the amount of time spent in an adverse health state become strong drivers of population ill-health. Simple phobia, social phobia, depression, dysthymia and alcohol dependence emerged as public health priorities.

Conclusions: From a clinical perspective, we tend to give priority to the disorders that exact a heavy toll on individuals. This puts the spotlight on disorders such as bipolar disorder and schizophrenia. However, from a public health perspective, disorders such as simple phobia, social phobia, depression and dysthymia—which are highly prevalent and tend to run a chronic course—are identified as leading causes of population ill-health, and thus, emerge as public health priorities.

Introduction

Mental disorders are gradually moving up to the higher levels in the hierarchy of leading causes of population ill-health (Murray and Lopez 1996; Menken et al 2000; Mathers and Loncar 2006; Begg et al 2008). For example, a study looking at the main contributors of years lived with disability found that the top-5 was composed of mental disorders, neurological and sense organ disorders, chronic respiratory diseases, diabetes mellitus and cardiovascular disease (Begg et al 2008). Nevertheless, many countries are still dedicating only a small fraction of their health care budget to mental disorders (World Health Organization 2006; 2008). This is unlikely to be a rational manner of resource allocation. Continuation of these policies is expected to increase rather than bridge the mental health gap (World Health Organization 2006; Lancet Global Mental Health Group 2007), since today's societies, with mostly knowledge-, service- and innovation-driven economies, are exerting pressure on the cognitive, social and creative skills of people, and these pressures are unlikely to decrease in the near future. At the same time, the current economies would benefit from a mentally fit and resilient population successfully engaged in today's economies. These important issues should place burden of disease studies with a focus on mental health in the limelight of attention of policymakers across sectors such as health, social affairs, employment and education. Our study differs from previous burden of disease studies in several ways. We estimate the burden of disease stemming from 15 DSM-III-R mental disorders and express burden of disease in terms of non-fatal disease burden, i.e. the number of years lived with disability (YLD). The YLD estimates are based on a representative population-based psychiatric survey in the Netherlands, allowing us to correct the YLD estimate for one disorder for existing comorbidities with other mental disorders or somatic illnesses. Previous burden of disease studies were often based on the assessment of medical experts, whereas in our research, we base our results on how the general population values the health states, in line with the Global Burden of Disease Study 2010 (Salomon et al 2012). Furthermore, burden of disease will be described at both the individual (per person) level and population level. It will be shown that ranking the disorders by disability leads to different hierarchies when taking the clinical perspective on individuals or a public health perspective on populations. A clear distinction between both perspectives is important to accurately inform decision makers, clinicians and researchers in the health care sector such that the efforts to scale up and improve the mental health services can be undertaken in a substantiated way.

Materials and methods

Sample

We used the data of the Netherlands Mental Health Survey and Incidence Study (NEMESIS). This study has been described in detail elsewhere (Bijl et al 1998). In brief, a random, stratified, multistage sample was obtained in three steps in 1996. First, municipalities were stratified by urbanisation, and 90 municipalities were drawn randomly and proportionately from these strata. Second, within each municipality, households were randomly drawn from the postal register. Finally, within each household, the person with the most recent birthday was selected on condition that he or she was aged between 18 and 65 years and was sufficiently fluent in Dutch to be interviewed. Eligible persons who were not immediately available were contacted later in the year. The response rate was 69.7%, resulting in a sample of 7,076 people. For 20 people (0.3%), the disability weight could not be computed and the effective sample size was, therefore, 7,056. The sample followed the same multivariate distribution over age, gender, civil status and urbanisation as the general Dutch population; however, males in the age group of 18–24 years were slightly underrepresented.

Measures

Demographics include gender, age, partner status, level of education, and employment status (working at least 8 h per week in a paid job). Mental disorders were assessed with the Dutch 1.1 version of the Composite International Diagnostic Interview (CIDI), which was developed by the World Health Organization for use by trained interviewers who are not clinicians (World Health Organization 1997; Ter Smitten et al 1998). The CIDI is a computerized psychiatric interview and generates 1-month, 1-year and lifetime prevalence rates of the DSM-III-R axis-I mental disorders. WHO field trials have documented acceptable reliability and diagnostic validity (Wittchen 1994). To be better able to study comorbidity and its effects on the disease burden, we used the CIDI without imposing the rules for the hierarchy among the disorders, meaning that if a person manifests with two disorders, for example, both a depression and schizophrenia, we count this as two distinct disorders, and not as a single disorder (a depression as part of schizophrenia). Somatic illnesses were self-reported by the respondents. Examples included chronic obstructive lung disease, emphysema, osteoarthritis, heart disease, diabetes mellitus. In total, the list contained 31 medical conditions. For each of the illnesses, it was also asked whether the patient received medication or another form of regular medical attention.

The self-reported illnesses were deemed to be measured with a greater reliability when the illnesses were said to be under medical attention. Health-Related Quality of Life (HRQoL) valuations were obtained from the Medical Outcome Study Short Form 6 Dimensions (SF-6D), based on the SF-36, using Brazier's algorithm (Brazier et al 2002). The SF-6D is a much used and well-validated instrument derived from the Medical Outcomes Study Short Form (Ware and Sherbourne 1992), and the one well-validated algorithm that was applicable to our data. It is of note that the SF-6D can describe as many as 18,000 health states, all referring to descriptive health states [i.e. all the permutations of the items (1) physical functioning, (2) role limitations, (3) social functioning, (4) pain, (5) mental health, and (6) vitality, each of which has five or six possible answers]. To obtain HRQoL valuations of all 18,000 health states would be a daunting task. Therefore, Brazier and colleagues used a sub-sample of 249 health states to elicit valuations in a representative sample (N = 836) from the general public in the UK. During a personal interview each respondent was asked to value the selected health states, and valuation was carried out using the standard gamble method, which was originally developed by Von Neumann and Morgenstern in 1953 (Von Neumann and Morgenstern 1953). In standard gamble, individuals are asked to choose between the certainty of living in a health state versus a treatment, which entails a chance of getting well at probability P and dying at probability 1-P. The idea is that people are more willing to accept a risky treatment that involves a higher risk of dying when their health-related quality of life is poor. The trade-off between quality of life and survival can be converted into disability weights (DWs) for each of the health states on a 0–1 scale, where 0 is the best possible health state (no disability) and 1 is the worst possible health state, equal to death (Gafni 1994). Finally, Brazier and colleagues used the health state valuations thus obtained in an econometric model to predict the values of all 18,000 health states that can be described by the SF-6D. The Brazier algorithm is based on this econometric model and can be obtained from John Brazier at the Sheffield University. We employed the algorithm developed for use in the SPSS statistical package.

Analysis

We linked disability weights to each of the DSM-III-R axis-I disorders. Here, we took two approaches: one without, and another with adjustments for comorbid mental disorders and somatic illnesses. Unadjusted DWs for each of the disorders were computed as the average DW of all respondents meeting the diagnostic criteria for a particular disorder. While this approach may portray a realistic picture of the disease burden in groups that meet the diagnostic criteria for a disorder, it can be criticized

for overestimating the disease burden attributable to a specific disorder when there are comorbid conditions that lend extra weight to the disability. There are various ways to adjust DWs for comorbidity (Andrews et al 1998). One is to select 'pure' cases presenting with a single disorder and average their DWs, but it should be mentioned that this approach has been criticized for being unrealistic, because comorbidity is the rule rather than the exception, and DWs would then be based on atypical disease patterns. A second approach is to attribute the disease burden to the hierarchically more important 'primary' diagnosis in the presence of 'secondary' disorders. However, assessing which disorders rank first and second is a difficult task, involving arbitrary decisions, and we preferred to stay clear of such complexities. Therefore, we took a third approach, adjusting for comorbidity by regressing DWs simultaneously on all the mental disorders and somatic illnesses. The regression coefficients are then interpreted as the DW of one disorder adjusted for the other disorders and illnesses in the model, thus resulting in adjusted DW estimates. The intercept (constant) in the regression model is 0.14 DW (with a standard error of 0.0015). The intercept represents the average DW, conditional on DSM-III-R disorders and illnesses. The intercept can be interpreted as the HRQoL decrement, attributable to unobserved factors affecting HRQoL such as minor illnesses, accidents and conditions that were not measured such as personality disorders. It should be noted that when a person presents with depression, he has the adjusted DW corresponding to depression (0.11) plus the base-rate DW (0.14), thus, a total DW equal to 0.25. In this way, the adjusted DWs were computed for all disorders. Finally, YLD were computed as the DWs attributable to a specific disorder multiplied by the number of people suffering from that disorder, while taking the duration of the disorder into account. This number of people starting to suffer from a disorder multiplied by the duration of the disorder is commonly referred to as person-years, which is approximated by the point-prevalence of a disorder. Finally, to facilitate extrapolation of our results to other countries, we calculated YLD per one million population: YLD/mln = DW * pyrs/mln. For data-analytical purposes, we used the 1-month prevalence rates as the best-available proxy for the point prevalence of the mental disorders.

To account for non-response, post-stratification weights were used in all analyses; see Bijl et al. (1998) for technical details. After weighting, the sample followed exactly the same multivariate distribution over age, gender, civil status and urbanisation as the population according to Statistics Netherlands (downloadable from www.cbs.nl). The Brazier algorithm for the SF-6D was executed in SPSS (version 15.0 for Windows), while all other analyses were conducted in Stata (version 8.2 for Windows).

Since the data were weighted, we used robust statistical techniques (the Huber-White Sandwich Method) to obtain correct sample errors (SE) and P values under weighting (Hox 2010).

Results

Characteristics of the studied population

The whole sample has, on average, a disability weight of 0.167 (SE = 0.001) on a scale of 0 (no disability; optimal health) to 1 (completely disabled; extremely poor health) indicating that people generally do not experience quality of life to its optimum, but 16.7% below its optimum (**table 1.1**). For each of the groups, the average unadjusted disability weight was calculated and t-tests were conducted to see whether the DWs differed significantly across categories. Differences between SF-6D based DWs are said to be clinically important when exceeding 0.041 (Walters and Brazier 2005). On average, women experience poorer quality of life than men. A greater amount of disability is experienced by people who had received less education, were not living with a partner, were jobless, and suffered from a mental disorder or a somatic illness. The mean age was 39.6 years (SE = 0.196; range 18–65 years) and age was significantly associated with disability: for every additional year of age, the DW showed a linear increase by a factor of 0.00049 (SE = 0.00014; t = 3.58; P<0.001), indicating that the quality of life decreases somewhat with increasing age. We found no evidence for a non-linear (quadratic) relationship between DW and age.

Table 1.1. Description of the sample: prevalence rates and standard errors (s.e.), disability weight (DW), and the p-value for the difference in the disability weights (N=7,056)

Variable	Prevalence	s.e.	DW	s.e.	P*
Male / Female	50.3% / 49.7%	0.76	0.159 / 0.193	0.002 / 0.002	<0.001
With / Without partner	70.0% / 30.0%	0.60	0.172 / 0.188	0.002 / 0.003	<0.001
High / Low education	60.8% / 39.2%	0.73	0.170 / 0.184	0.002 / 0.003	<0.001
With / Without a job	70.7% / 29.3%	0.68	0.163 / 0.205	0.002 / 0.003	0.004
Without / With disorder	85.6% / 16.4%	0.48	0.164 / 0.242	0.001 / 0.004	<0.001
Without / With illness	60.7% / 39.3%	0.71	0.149 / 0.220	0.002 / 0.002	0.003
* P-value for the DW difference relative to the first (reference) category of the variable.					

Unadjusted disability weights

Table 1.2 presents the unadjusted disability weights, the number of person years spent in illness per million population (pyrs/mln), and the number of years lived with disability per one million (YLD/mln) for each of the 15 DSM-III-R axis-I disorders. It is of note that people who meet the diagnostic criteria for any mental disorder have a

health-related quality of life that is 24% lower than the theoretical maximum. Mood disorders, panic disorder, obsessive compulsive disorder, social phobia, eating disorders and schizophrenia emerged as the disorders with a relatively high disease burden at individual level: these disorders are associated with a relatively large DW.

Table 1.2. Unadjusted disability weights (DW), standard errors (s.e.), person years (pyrs) spent in illness per annum, and years lived with disability (YLD) per 1 million population by DSM-III-R mental disorder, not adjusted for co-morbidity and somatic illnesses (N=7,056)

Disorder	DW	s.e.	pyrs/mln	s.e.	YLD/mln	s.e.*	Ranking
Any disorder	0.24	0.004	164,456	4,829	39,469	1,333	-
Mood disorders	0.34	0.008	39,329	2,380	13,508	868	ii
. depression	0.35	0.009	26,037	1,969	9,117	728	3
. bipolar	0.31	0.018	6,128	957	1,925	317	10
. dysthymia	0.36	0.012	16,854	1,524	6,007	585	4
Anxiety disorders	0.26	0.005	96,741	3,685	25,275	1,073	i
. panic	0.34	0.013	14,654	1,466	5,027	534	6
. agora	0.30	0.014	10,155	1,194	3,078	385	8
. social	0.28	0.008	37,207	2,304	10,305	710	2
. simple	0.30	0.007	55,437	2,895	16,804	951	1
. generalised	0.25	0.020	8,147	1,141	2,066	329	9
. obsess. comp.	0.33	0.034	2,661	606	880	219	12
Substance use	0.20	0.006	58,067	3,187	11,553	726	iii
. alcohol abuse	0.16	0.007	24,821	2,335	4,057	412	7
. alc. dependence	0.21	0.010	27,145	2,248	5,744	545	5
. drugs abuse	0.22	0.027	2,595	697	560	169	14
. drugs dependence	0.22	0.021	7,061	1,214	1,524	305	11
Eating disorders	0.32	0.030	2,524	576	810	199	13
Schizophrenia	0.30	0.056	1,598	483	483	170	15

* Standard error calculated by using the standard rules when multiplying two variables, under the assumption that both (DW and pyrs/mln) are independent.

Unadjusted years lived with disability

At population level, both the number of affected people and the actual disease duration become important drivers of disease burden. Both factors are captured in the number of person-years spent in illness. It appeared that the common mental disorders, such as anxiety and substance use disorders, emerged as prominent causes of disease burden. It is of note that simple phobia, a seemingly mild disorder, is not only associated with a relatively large number of person-years, but has in addition a markedly high average disability weight, making it the single leading cause of disability in the field of mental health in the Netherlands. Here, we must emphasize that the YLD have not yet been adjusted for comorbidity.

Adjusted disability weights
Table 1.3 presents the results when statistical adjustments are made for all
comorbidities including somatic illnesses, but not for demographics such as age and
gender. It appears that the adjusted DWs are lower than the unadjusted ones by 28%
on average. A large reduction in the DWs was to be expected, because now we assess
the unique contribution of each of the distinct disorders to disease burden. The DWs
of eating disorders and simple phobias are particularly reduced after adjustment for
comorbidity; somatic illnesses and drug dependence are least affected by
adjustment.

Table 1.3. Adjusted disability weights (adj DW), standard error (s.e.), person years (pyrs) spent in
illness per annum, and years lived with disability (YLD) per population of 1 million by DSM-III-R
mental disorder (N=7,056)

Disorder	adj DW	s.e.	pyrs/mln	s.e.	YLD/mln	s.e.[*]	Ranking
Mood disorders	0.27	0.002	39,329	2,380	10,790	658	ii
. depression	0.25	0.003	26,037	1,969	6,524	500	3
. bipolar	0.24	0.004	6,128	957	1,457	229	10
. dysthymia	0.22	0.003	16,854	1,524	3,684	337	5
Anxiety disorders	0.20	0.001	96,741	3,685	19,045	732	i
. panic	0.21	0.003	14,654	1,466	3,117	315	7
. agora	0.20	0.002	10,155	1,194	2,060	243	8
. social	0.18	0.001	37,207	2,304	6,580	409	2
. simple	0.17	0.001	55,437	2,895	9,349	491	1
. generalised	0.19	0.004	8,147	1,141	1,585	224	9
. obs.comp.	0.23	0.006	2,661	606	625	143	12
Substance use	0.15	0.001	58,067	3,187	8,904	492	iii
. alc. abuse	0.13	0.001	24,821	2,335	3,314	313	6
. alc. dependence	0.16	0.001	27,145	2,248	4,268	354	4
. drugs abuse	0.16	0.004	2,595	697	424	114	13
. drugs depend.	0.19	0.004	7,061	1,214	1,366	237	11
Eating disorders	0.16	0.007	2,524	576	404	94	14
Schizophrenia	0.23	0.001	1,598	483	363	110	15
Somatic illness	0.20	0.003	392,600	6,132	80,203	1737	-

[*] *Standard error calculated by using the standard rules when multiplying two variables, under
the assumption that both (adj DW and pyrs/mln) are independent.*

Adjusted number of years lived with disability
At population level, the YLD are on average 28% lower after adjustment, but the rank-
order as causes of population ill-health remains unaffected; both the unadjusted and
adjusted YLD show the same hierarchy: simple phobia emerges as the leading cause
of non-fatal disability, followed by social phobia, depression, dysthymia and alcohol
dependence and these are followed by eating disorders and schizophrenia (mainly
due to the small number of affected people).

Discussion

Main findings

Mental disorders are a major cause of disease burden, but a distinction should be made between disease burden at individual level and at population level. At individual level, depression, bipolar disorder, dysthymia, panic disorder, obsessive compulsive disorder and schizophrenia have a severe and adverse impact on health-related quality of life. In addition, people presenting with eating disorders and simple phobias tend to experience greatly reduced health-related quality of life, but this is largely attributable to comorbid conditions. At population level, the number of people and the amount of time spent in an adverse health state become prominent drivers of population ill-health. Leading causes of years lived with disability are simple phobia, social phobia, depression, dysthymia and alcohol dependence — all of which are highly prevalent disorders or disorders that tend to persist over time. Our top-5 list remains unaltered after adjusting for comorbidity, indicating that each of these disorders significantly contributes to population ill-health in their own right, even when some of these conditions are often accompanied by concurrent disorders, as is the case for the simple phobias, social phobia and dysthymia. It may, therefore, be assumed that these disorders can be seen as priority targets for further scientific scrutiny and public health intervention.

Context and other studies

We need to place our findings in the wider context of the literature. The simple phobias are the most common mental disorders, and in our study, the phobias dominate other disorders in terms of their contribution to population ill-health. Previous burden of disease studies (Murray and Lopez 1996; Menken et al 2000; Mathers and Loncar 2006; Begg et al 2008), while using different methods to estimate disability weights, did not specifically look into the disability caused by simple phobias.

A cost of illness study (Smit et al 2006a) shows that the simple phobias are associated with very modest health care costs, indicating that people with simple phobia receive relatively little medical attention. Our data suggest that this might have important implications. First, many people suffer from these phobias and the phobias are often precursors of other mental disorders (de Graaf et al 2003). Treating the phobias may, thus, have preventive value. Second, the corresponding disability weight is relatively high: 0.30 (unadjusted) and 0.17 (adjusted). The fact that the adjusted disability weights are quite lower indicates that simple phobias are often

comorbid with other conditions, but have, nevertheless, a fairly large disability weight in their own right. It is worth noting that our results are consistent with the literature. First, in one of our own studies (Smit et al 2006a), we demonstrated that the simple phobias are associated with remarkably high economic costs due to absenteeism from work, but not due to health care uptake. This lends support to the impression that people with simple phobias are not often the recipients of professional care, and yet often stay away from their work, suggesting a certain amount of functional impairment stemming from this condition. The latter is also observed in the European-wide ESEMeD study: the simple phobias do not attract much medical attention, but are notoriously associated with work absenteeism (ESEMeD 2004). A recent study on the effect of mental disorders on productivity shows that simple phobia is associated with additional days of absenteeism, although the effect is less pronounced (de Graaf et al 2012a).

Social phobia is a prevalent disorder with an onset typically in adolescence or early adulthood, and it is likely to run a chronic course when left untreated. These factors combine to make social phobia a disorder associated with a sizeable number of person-years spent in illness: 37,207 pyrs/mln in our study, ranking second after the simple phobias. In addition, social phobia is likely to have an adverse impact on academic achievement and professional performance, especially in the context of today's communication and service-driven economy (Weehuizen 2008). Its relatively large disability weight of 0.28 (unadjusted) and 0.18 (adjusted) has to be placed in this context. Social phobia featured prominently in the top-5 of leading causes of disability in the HRQoL study by Saarni et al. (2007) and our evidence adds to theirs. The presence of major depression on our top-5 list does not come as a surprise, because it has been consistently identified as a leading cause of disability (Murray and Lopez 1996; Bowie et al 1997; Mathers et al 2001; Mathers and Loncar 2006; Saarni et al 2007). Currently, depression is the single leading cause of non-fatal disease burden in high-income countries and it is projected to become the second leading cause of disability-adjusted life year (DALY) disease burden (which also accounts for mortality) by 2020, second only to ischemic heart disease (Murray and Lopez 1996). More recent projections predict that depression might become the single leading cause of DALY disease burden in the high-income countries by the year 2030 (Mathers and Loncar 2006). Dysthymia has been under-studied as a contributor to population ill-health, possibly because its disability weight was often assumed to be equivalent to mild depression, and because major depression and dysthymia were sometimes combined into a single disease category (Vos and Mathers 2000). However, more recently, Saarni et al. (2007) identified dysthymia as the single leading cause of health-related quality of life decrement in Finland when basing the

disability weight of dysthymia on valuations directly obtained from the target population. Saarni argued that the general population tends to weigh health states more heavily when these persist over time. Indeed, we obtained a DW for dysthymia of 0.36, adjusted to 0.22 when comorbidity was taken into account, which is, methodological differences aside, higher than the weight of 0.14 for mild depression (Stouthard et al 1997) commonly used as a proxy for dysthymia in other studies. This would explain why dysthymia has not often been identified as a significant cause of disability in a population. Elsewhere, we showed that dysthymia is associated with the highest economic costs among the common mental disorders, thus, underscoring its relative importance as a disabling condition (Smit et al 2006a).

Alcohol dependence also ranks in the top-5 of our list of most disabling disorders at population level. There is no doubt about its importance as a risk factor for population ill-health and our study is more likely to underestimate than to over-estimate its importance. After all, it is a condition associated with premature death (Gmel et al 2003), and we did not take mortality into account. Moreover, alcohol dependence is not only a clinical endpoint, but perhaps more importantly a risk factor for more than 60 medical conditions, ranging from cancer to liver sclerosis with all due consequences (Rehm et al 2003; Chisholm et al 2004b; Konnopka and König 2009; Pillai et al 2012). Alcohol dependence is, therefore, an important public health target and one may speculate that this is also true for the less severe but more common manifestations of alcohol misuse, such as hazardous and harmful use that are not part of the DSM classification.

Strengths and limitations
The strengths of this study include the population-based large-scale representative dataset on which the analyses were based and the use of formal DSM classifications of the mental disorders which were reliably assessed with the CIDI. The use of national, rather than regional or global, burden estimates makes our findings particularly useful to the Netherlands' health service system.

Another strength is the disability weights being derived from the general population. There are several ways of eliciting HRQoL valuations (e.g. from professionals in the medical field), but ultimately we need to understand how populations evaluate their own health-related quality of life (Saarni et al 2007).

A final strength is that we could compute both unadjusted and adjusted years lived with disability (YLD) either by accepting comorbidity with other mental disorders and somatic illnesses as a fact of life, or adjusting for comorbidity. Unadjusted YLD portray an accurate picture of the burden of disease in groups of

people that are likely to have comorbid conditions—after all, in real life, we do not encounter people who have been adjusted for comorbidity. Therefore, unadjusted YLD may have value from a public health perspective. However, when the aim is to assess the YLD attributable to a specific disease, then adjusted estimates are preferred, because adjusted YLD give information about the amount of disability due to a specific disorder without confounding by co-occurring conditions. In our study, we explored both approaches and were, thus, able to shed light on both issues.

We acknowledge the following limitations of our study. First, people with severe conditions may have been unable to participate in this population-based survey because they were hospitalised, and this is likely to have resulted in an under-estimation of the disease burden. We expect the DW estimates of schizophrenia, bipolar disorder, severe depressions, and the more severe cases of anorexia nervosa, to be most affected by this.

Second, YLD captures only the non-fatal component of disease burden. As we did not include years of life lost (YLL) due to mortality, this results in an under-estimation of the overall disease burden, in particular for conditions associated with excess mortality such as some of the somatic illnesses and the substance use disorders (Murray and Lopez 1996).

Third, we used the Brazier algorithm, which provides estimates of utility, to calculate disability weights, by converting the utility weights into disability weights using the formula *disability weight = 1 - utility*. However, utilities and disability weights are not exact complements (utilities can assume values less than 0), so applying this straightforward translation could bias our results. Also, the Brazier algorithm was based on assessments in a sample of British people, while our sample was from the Netherlands. This may have distorted our outcomes somewhat, although it is unlikely to change the overall results in a substantial way, as differences in Western Europe between national value sets, such as the set for the SF-6D, are small (Craig et al 2009). The use of the Brazier algorithm also limits the comparability with burden of disease studies using different methods for estimating health-related quality of life (e.g. AQoL, EQ-5D, HUI3, 15D). Hawthorne et al (2001) demonstrate the different outcomes associated with these different instruments and state that no single instrument can claim to be the 'gold standard', and that differences in outcomes are in part attributable to differences in coverage of the different dimensions of health-related quality of life (Hawthorne et al 2001).

Fourth, available data made us use the 1-year prevalence rates of the somatic illnesses, while we relied on the 1-month prevalence rates of mental disorders. Thus, we could have over-estimated the YLD due to the somatic illnesses.

However, the somatic illnesses were all chronic, making it unlikely that 1- and 12-month prevalence rates would differ drastically. We therefore expect that this did not substantially impact on our findings.

An additional limitation is that data on somatic illnesses is based on self-report rather than diagnostic assessment, as well as the fact that not every somatic illness was adjusted for in the model. The fact that we did not adjust for every illness, but only for 31 illnesses, may have inflated DW estimates. In addition, since data on self-reported illnesses not under medical attention were excluded, further DW inflation might have occurred. While we feel that the people should be the ultimate judges of their own health, a panel of lays may be associated with limitations that are worth noting, such as lesser consistency, and the possibility that (healthy) lays have difficulties passing judgments on the severe conditions. This may have caused some under-estimation of the disability weights associated with the more severe disorders. Indeed, regarding the severe conditions, Brazier et al. (2002), pointed out that ''inconsistent estimates and over prediction of the value of the poorest health states'' might be seen as a limitation of their method.

A final limitation is that our data were based on DSM-III-R criteria, whereas one would prefer DSM-IV criteria or, in the future, DSM-5 criteria.

Implications

Our study showed that a clinical perspective on individual disease burden or a public health perspective on the disease burden at population level result in different health priorities. After all, the clinical approach brings individual suffering into focus, but at population level, the number of affected people and the duration of disorders become key drivers of population ill-health.

The fact that different perspectives lead to different conclusions about health priorities may not come as a complete surprise, but the difference can be quite striking. To illustrate, at individual level, one could have the impression that the simple phobia is not a priority for intervention, but the sheer number of years lived with disability due to simple phobia in the population does raise questions about the ways to alleviate its disease burden. Such paradoxes can be confusing. In debates about priorities in health care, it is important to understand how these paradoxes can arise—in particular when both clinicians and public health decision-makers are at the same table. Our study suggests that the clinicians may need to reconsider the population-level impacts of disorders like simple phobia, social phobia and dysthymia in terms of HRQoL—especially in relatively disadvantaged groups such as women, jobless people and people with lower attained levels of education, those without a

partner and those presenting with comorbid conditions—and perhaps identify phobias as targets of treatment more frequently. Conversely, decision-makers in the field of public health may have to be persuaded that for these disorders, clinical interventions may perhaps not be appropriate, but that attention should be directed to the question whether acceptable, effective and economically affordable (minimally supported) self-management interventions for these disorders can be developed (for example, as self-help interventions offered over the internet), evaluated and implemented on a scale proportional to their disease burden. Given the relatively low burden on individual level, in combination with the large burden on population-level, self-help interventions, when deemed appropriate, could play an important role in bridging the health gap for this population.

Part II

Established methods for assessing the cost-effectiveness of interventions

The role of support in web-based treatment for depression and anxiety in adults: cost-utility analysis alongside a randomized trial

In preparation

Authors: Kleiboer A, Lokkerbol J, Donker T, van Straten A, Mihalopoulos C, Cuijpers P, Smit F

Abstract

Background: Economic and demographic developments highlight the need to identify interventions capable of reducing the disease burden of people with depression and anxiety while putting minimal pressure on financial and human resources.

Objectives: To assess the cost–utility of a web-based intervention based on Problem-Solving-Treatment (PST) with varying levels of coaching aimed at treating mild to moderate depression and anxiety in adults versus care-as-usual (CAU).

Method: A randomized controlled trial with four arms was used to compare an eHealth intervention with different levels of support (no support, support on request, weekly support) with a control condition of weekly non-specific support, either by chat or email. Cost-utility was assessed over 12-months to identify the intervention associated with the least incremental cost per additional QALY gained.

Results: At 12 months, the incremental cost in comparison to the control condition was -€991 (SD 1,266) in the intervention without support, €469 (SD 1,352) with support on request, and -€225 (SD 1,341) with weekly support. Incremental QALYs gained were 0.009 (SD 0.023), 0.018 (SD 0.022) and 0.003 (SD 0.023), all non-significant. The intervention conditions with no support and weekly support achieved (non-significantly) more health effects for on average lower costs than the control condition. The intervention conditions with no support, support on request and weekly support had a 79%, 47% and 56% probability of being more cost-effective than the control condition at the €20,000 willingness-to-pay (WTP) threshold value. The intervention condition with support on request was associated with a cost of €25,697 per QALY gained. The intervention conditions did not show greater improvements on CES-D (t=1.532, df=215, p=0.937; t=0.009, df=216, p=0.504; t=0.342, df=214, p=0.634, respectively) and HADS-A (t=1.499, df=215, p=0.933; t=-0.513, df=216, p=0.305; t=0.928, df=214, p=0.823, respectively).

Conclusion: The PST intervention without support has a 70%-80% probability of being more favourable than weekly supportive chats or emails in terms of cost-utility over the full range of considered WTP values. The fact that the intervention groups that received weekly support or support on request did not dominate in terms of cost-utility indicates that more therapists' involvement does not necessarily lead to a more favourable incremental cost per QALY. We therefore conclude that less therapists' involvement is not only associated with a more favourable cost-utility, but may in addition amount to treating more patients without unduly exerting pressure on available human resource capacity. More research is needed to see whether high

dropout rates biased our results and whether improved adherence increases the impact of the web-based intervention.

Trial registration: Netherlands Trial Register, TC1355.

Introduction

Depressive and anxiety disorders are highly prevalent and are associated with substantial disease burden (Lokkerbol et al 2013) and economic costs stemming from health care uptake and productivity losses (Smit et al 2006a). Even with optimal coverage of evidence-based treatment, it is expected that a substantial part of this disease burden is not averted (Andrews et al 2004), emphasizing the importance of investigating new ways to deliver treatment. Internet-based self-help interventions based on evidence-based treatments have shown to be effective in reducing symptoms in people with mild to moderate anxiety and depression (Spek et al 2007a; Cuijpers et al 2010; Haug et al 2012). However, it is not clear how these interventions should be implemented in practice. One important issue in this respect is whether they should be delivered with human support (coaching) or as 'pure' self-help. Meta-analytic studies have shown that Internet interventions that are offered with support have higher effect sizes than interventions that are delivered without support (Spek et al 2007a; Richards & Richardson 2012). However, a system without human support is much easier and cheaper to implement than a system where patients are guided by a coach, as this would put pressure on therapist time. Therefore, it is important to know whether Internet-based interventions delivered without professional support are more cost-effective than Internet-based interventions with support; are the benefits of additional coaching "worth" the extra costs?

This paper reports the results of an economic evaluation that was conducted alongside a randomized controlled trial comparing an online intervention for reducing symptoms of anxiety or depression with varying levels of coaching support compared to non-specific support in the control condition.

Methods

Design
The study design was described in detail elsewhere (Donker et al 2009). In brief, an economic evaluation was conducted alongside a pragmatic randomized trial with three intervention arms and two control arms. Measurements were conducted at baseline (t_0), posttest at six weeks after baseline (t_1) and follow-up at three and twelve months after baseline (t_2 and t_3). Randomization was conducted after baseline at the individual level using block randomization with variable block sizes.

The allocation schedule was derived by computer using a random number generator and was conducted by an independent researcher. With 100 participants per arm, the trial was powered to detect a standardized effect size (Cohen's d) of 0.35 (or larger) in a one-tailed test at $\alpha=0.05$ (1-sided) with a power of $(1-\beta)=0.80$. Smaller effect sizes were not considered to be clinically important. It is worth noting that the trial was powered for testing hypotheses about differences in clinical outcomes (treatment response), but not to test economic hypotheses. For the health economic evaluation a probabilistic medical decision-making approach is used instead.

Trial results are reported in agreement with the CONSORT statement (Schulz et al 2010) and the economic evaluation with the CHEERS statement (Husereau et al 2013). The study was approved by the Medical Ethics Committee of the VU University Medical Centre (ref: 2008-11) and was registered with the Netherlands Trial Register (NTR; ref: TC1355). The economic evaluation used a Dutch health sector perspective to which the costs were added of productivity losses stemming from absenteeism and lesser efficiency while at work (presenteeism).

Participants
Participants (aged 18 years or older) who met the inclusion criteria of a score of 16 or higher but less than 40 on the Center for Epidemiologic Studies Depression scale (CES-D) and/or 8 or higher and less than 15 on the anxiety subscale of the Hospital Anxiety and Depression Scale (HADS-A) were considered to have mild to moderate symptoms of depression or anxiety. Participants were excluded if they: 1) had insufficient knowledge of the Dutch language; 2) had no access to a computer with a fast internet connection; 3) reported active suicidal plans (based on a self-report screening question (SQ) developed by Marks and colleagues (Gega et al 2005)); and, 4) received treatment by a mental health specialist (i.e. a psychologist or a psychiatrist) at the time of recruitment. Participants were allowed the use of prescribed medication for anxiety and depression disorders with stable dosage (for at least one month prior to assessing eligibility).

Recruitment
Participants were recruited by placing banners on websites and advertisements in national and local newspapers. As shown in **figure 2.1**, a total of 1,319 potential participants were assessed with respect to the inclusion criteria. Of these, 782 were excluded, with 537 (69%) not meeting inclusion criteria, 140 (18%) not completing the screening, 104 (13%) declining participation, while 1 person was excluded due to

a logistic error. The remaining 537 participants were randomized to one of five arms of the trial:

1. web-based problem-solving therapy (PST) without support (n=107),
2. web-based PST with support on request (n=108),
3. web-based PST with weekly support (n=106),
4. a control condition consisting of non-specific coaching either by email or chat (n=110), or
5. a second control condition consisting of a waitlist (n=106).

Since the economic questionnaire was not presented to participants in the waitlist control condition, as this was not considered during the original design of the trial, this cost-utility analysis only considers the first four arms, where the fourth arm, non-specific email or chat support, serves as control condition. Non-specific email or chat support was offered to control for the non-specific effects of coaching. Participants had access to usual care in all conditions.

Figure 2.1. Flowchart of participants in the trial

Enrollment

Assessed for eligibility (n=1,319)

Excluded (n=782)
- Not meeting inclusion criteria (n=537)
- Declined participation (n=104)
- Screening not completed (n=140)
- Other (n=1)

Randomised (n=537)

Allocation

| Intervention *without* support (n=107) | Intervention with support *on request* (n=108) | Intervention with *weekly* support (n=106) | Non-specific email or chat support (n=110) | Waitlist (n=106) |
| Received allocated intervention (n=83) | Received allocated intervention (n=95) | Received allocated intervention (n=89) | Received allocated intervention (n=98) | |

Follow-Up

| Lost to follow-up / discontinued intervention 69% | Lost to follow-up / discontinued intervention 54% | Lost to follow-up / discontinued intervention 60% | Lost to follow-up / discontinued intervention 45% |

Analysis

| Analysed (n=107) | Analysed (n=108) | Analysed (n=106) | Analysed (n=110) |

Intervention: web-based problem solving treatment

Participants in the treatment conditions received the web-based intervention 'Alles onder controle' either without coaching, with coaching on request, or with pre-specified weekly coaching. The intervention is a brief self-help treatment based on problem solving techniques (Bowman et al 1995; 1997). The intervention takes five weeks to complete and consists of one lesson per week. Each lesson consists of information, examples and exercises. In the first lesson, respondents describe what they think is important in their lives, they make a list of their problems and concerns, and divide these into three categories: unimportant problems (problems unrelated to things that matter to them), solvable problems, and problems that cannot be solved (e.g. loss of a loved one). In the following lessons, a different strategy is proposed for each type of problem, to either solve or deal with the problem. The general idea of the intervention is that participants learn to regain control over their

problems. The intervention has shown to be effective in reducing symptoms of anxiety and depression in two previous randomized controlled trials (Warmerdam et al 2008; Van Straten et al 2008) when delivered online with weekly support.

Participants receiving the web-based intervention without support went through the program by themselves without being assigned a coach, receiving automated emails with general information regarding the purpose of the lessons as well as tips. Participants in the condition receiving support on request were assigned to a coach and had the option of contacting the coach via the program website in case of questions, or after completion of a lesson to ask for feedback. Participants in the condition receiving weekly support were actively approached by their coach after completion of a lesson, again via the program website, to provide feedback. In addition, participants in this condition had the option of contacting the coach via the program website if they had any questions.

Support was aimed to guide respondents through the self-help method and involved: (1) motivating the respondent to continue the program, (2) providing feedback to the exercises and explaining the PST techniques in more detail if needed, and (3) showing empathy. Coaches were Master level students in clinical psychology that were trained for approximately six hours to provide support. A coaching manual was provided and to ensure treatment fidelity all feedback was reviewed by a supervisor (AK or TD) before it was communicated via the program website. Coaching was either offered via email or using chat.

Respondents in all conditions received weekly reminder emails from their coach or from the research team if one was assigned to receive the web-based intervention without support. This has been recommended by previous studies to increase adherence (Nordin et al 2010).

Control condition: non-specific coaching
A coach was assigned to all participants in the chat or email condition either as weekly chat sessions or weekly email contact. The coach provided support based on non-directive conversation skill techniques such as Client-Centered Therapy (Rogers 1951), communication skills (Molen et al 1995) and clinical management used in the National Institute of Mental Health treatment for adolescents with depression study (TADS) (NIMH). The coach was allowed to give general support only, while at the same time avoiding specific techniques from other formal psychotherapeutic interventions. Coaches were again Master level students in clinical psychology

trained for approximately six hours. A coaching manual was provided and a supervisor supervised the chat sessions and reviewed all emails. Coaches who provided feedback to respondents in the chat or email condition were not the same coaches that provided feedback to the respondents assigned to follow the web-based intervention to prevent contamination.

Measures

Participants completed online self-report questionnaires at baseline and after 6 weeks, 3 months and 12 months. Primary outcomes were symptoms of depression and symptoms of anxiety using the CES-D and the HADS-A (for details see Donker et al 2009). This health economic evaluation reports clinical outcomes, but will mainly focus on (changes in) quality of life and economic costs.

Quality of life

The generic outcome measure of interest, quality of life, was assessed using the EuroQol-5D (EQ-5D). The EQ-5D consists of five items measured on a 1 to 3 scale, measuring respondents' mobility, self-care, pain, usual activity and psychological status (EuroQol Group 1990). For each item, a respondent can indicate that they experience no problem, moderate problems, or severe problems, resulting in 243 different health states. Each health state was valued using Dutch tariffs to obtain utilities (Lamers et al 2005). Utilities are a measure of the strength of preference for the various health states measured by the questionnaire. Utility values are anchored between 0 and 1 where 0 represents death and 1 represents full health. Multiplying the utility value by the length of time in that particular health state allows the calculation of Quality Adjusted Life Years (QALYs). QALYs are particularly important in economic evaluation as they allow comparison of the economic evaluations across different disorders and diseases.

Measuring resource use

Information on the participants' use of health services was obtained with the Netherlands Institute of Mental Health and Addiction (Trimbos institute) and Institute of Medical Technology Assessment Cost Questionnaire for Psychiatry (TIC–P; Hakkaart-van Roijen 2002). With this questionnaire, patients registered the number of general practice visits, sessions with psychiatrists, hospital days, etc. over the past four weeks at t_0, t_2 and t_3. In addition, the number of 'work loss' days (absenteeism from work) and the number of 'work cut-back' days (presenteeism; reduced efficiency at work while feeling ill) were also measured with the TIC–P. Questions were limited to the use of health services and productivity. Travel and

parking costs in the context of receiving health care were not part of the questionnaire but were estimated from the TIC-P's health care uptake data.

Cost of services

Costs were expressed in euro (€) for the reference year 2011 on a per participant basis for the period of one year, using linear interpolation to estimate the use of health services between the four-week periods that were assessed in the questionnaire. As the time frame of this study was restricted to one year, costs and effects were not discounted. All costs related to health care utilization are standard (full economic) cost prices as reported in the Dutch Costing Manual (Hakkaart-van Roijen et al 2010) and indexed to represent 2011 prices.

Screening costs were not considered, as they are equal across all conditions and therefore cancel out when comparing the cumulative costs of the intervention conditions and the control condition.

Intervention costs included hosting and maintenance costs of €44 (details can be obtained from the corresponding author) and costs related to time spent on sending reminder emails and coaching. The latter was calculated as the number of hours spent on sending reminders and coaching sessions multiplied by the hourly rate of €161 for a psychologist, because psychologists, not students, will act as coaches in the clinical setting. Time spent on reminder emails was estimated to be just under 5 minutes (0.08 hours) on average. Time for writing feedback in response to a completed lesson of the web-based intervention and the time for writing an email or a chat session was 25 minutes (0.42 hours) on average, which was constant over all conditions. The intervention costs for the web-based intervention without support are €44 + 1*0.08*€161 = €56.88, for the web-based intervention with support on request this equals €44 + 6*0.08*€161 + 0.5*0.42*€161 = €155.09, and for the web-based intervention with weekly support this equals €44 + 6*0.08*€161 + 3*0.42*€161 = €324.14. Costs in the control condition resulting from therapists' time when delivering non-specific chat or email support were disregarded in the analysis. Although the control condition controls for non-specific effects of coaching and is therefore expected to generate some effects, the relatively high use of health services to this end would not occur in real practice.

Direct medical costs are the costs of treatments offered by a range of health services (see **table 2.1**). These costs were computed by multiplying the number of health service units (consultations, hospital days, etc.) by their standard full economic cost price. To these we added the costs of medication, calculated as the cost price per standard daily dose (www.medicijnkosten.nl) which included the pharmacist's

claw-back rate as well as 6% VAT (which is non-deductible for Dutch citizens), multiplied by the number of prescription days, plus the pharmacist's dispensing costs of €6.00 per prescription.

Travel costs, which arise when patients travel to health service providers and pay for parking, were not collected via the questionnaire, but estimated at €0.21/km and €3.14/h parking time (**table 2.1**), based on average travel distance to health care services.

Table 2.1. Direct medical and direct non-medical costs by health service type

		Direct Medical Costs (in 2011 €)		Direct Non-Medical Costs (in 2011 €)
Health service type	Unit	Unit cost price [a]	Km	Unit cost price [b]
General practitioner	Contact	29.27	1.1	3.37
Company doctor	Contact	73.11	17.6	6.84
Social worker	Contact	67.95	5	4.19
Regional mental health service	Contact	178.78	10	5.24
Alcohol and drug consultant (CAD)	Contact	178.78	10	5.24
Independent psychotherapist	Contact	80.50	5	4.19
Independent psychiatrist	Contact	107.69	5	4.19
Physiotherapist	Contact	37.64	2.2	3.60
Medical specialist general hospital	Contact	75.28	7	4.61
Complementary and Alternative Medicine	Contact	53.00	5	4.19
Day treatment	Day	161.00	7	4.61

[a] Integral unit cost prices (cf. Hakkaart-van Roijen et al 2010).
[b] Costs = (0.21 * km) + 3.14, where €0.21 = cost per km; €3.14 = 1h parking time (cf. Hakkaart-van Roijen et al 2010).

Cost of production losses

Production costs due to absenteeism and "presenteeism" were also included in the current study. The costs of being absent from work were evaluated using the average age- and gender-specific hourly gross incomes obtained from Hakkaart-van Roijen et al. (2010), see **table 2.2**. Second, presenteeism costs, which are incurred when people are ill but continue to work with reduced efficiency, were estimated by multiplying the number of work cut-back days (i.e. the number of days actually worked when ill), by a self-reported inefficiency score, ranging between 0 and 1 (0, as efficient as when in good health; 1, totally inefficient). The resulting number of days of productivity loss was again valued using the age- and gender-specific hourly wages by Hakkaart-van Roijen et al. (2010), again inflated to represent 2011 prices.

Table 2.2. Productivity costs by gender and age class

	Productivity costs per hour (€, 2011 [a])	
Age Classes	Men	Women
15-19	*9.88*	*8.97*
20-24	*18.18*	*17.59*
25-29	*24.77*	*24.19*
30-34	*30.37*	*28.20*
35-39	*34.85*	*29.96*
40-44	*37.25*	*29.76*
45-49	*39.25*	*29.61*
50-54	*40.00*	*29.96*
55-59	*40.33*	*30.21*
60-64	*40.07*	*29.36*
65+[b]	*0.00*	*0.00*

[a] *Costs are indexed using the Collective Labour Agreement index rate.*
[b] *People aged 65+ are assumed to not have a paid job*
Source: Hakkaart-van Roijen et al. (2010).

Analysis

The analysis of the outcomes was conducted in accordance with the intention-to-treat principle. The expectation maximization (EM) imputation technique as implemented in SPSS version 22 (IBM Corp 2013) was used for handling loss to follow-up (Little and Rubin 1987). In the expectation maximization imputation model, baseline outcome measures, age, income, condition and working hours per week were used as predictors, because they were significant predictors of the outcome measures at follow-up.

Differences in CES-D and HADS-A outcomes between the intervention conditions and the control condition at 12-month follow up were tested using one-sided t-tests.

Utilities and costs were first corrected for baseline differences by multiplying individual outcomes within each condition with a factor that adjusted the group mean to the overall, grand mean. Adjusted baseline measurements were then used to adjust outcomes at later measurements by applying the same growth factors as observed between the unadjusted measurements for each participant. The measurements at baseline and follow-ups, now corrected for baseline differences, were then linearly interpolated to provide estimates throughout the full 12-month period. QALYs and cumulative costs were then calculated using the Area Under the Curve (AUC) method to use all available measurements.

As the trial is not powered to detect differences in QALYs and costs, the main focus is on reporting probabilistic outcomes, as is common when using cost-utility analysis within the medical decision-making framework.

Economic evaluation

The incremental cost-effectiveness ratio (ICER) was calculated as $(C_1-C_0)/(E_1-E_0)$, where C_1-C_0 is the difference in average annual per participant cost and E_1-E_0 is the difference in average QALY of participants in the experimental conditions relative to the control condition (subscripted 1 and 0 respectively). In other words, the incremental cost-effectiveness ratio is the difference of the cumulative costs between the conditions divided by the difference in QALY gains. The ICERs were evaluated by means of non-parametric bootstrapping (2500 iterations). These simulated ICERs are presented as a scatter over the cost-effectiveness plane (see **figure 2.2**), with differences in costs on the vertical axis and differences in QALYs on the horizontal axis. If the ICERs appear in the top left-hand quadrant of the plane, more costs are paid for lesser effectiveness; the intervention is then unacceptable from a cost-effectiveness perspective and conventional care remains the treatment of choice. If the ICERs appear in the lower right-hand quadrant, lesser costs (savings) are then associated with health gains; the intervention dominates and is acceptable from a cost-effectiveness point of view. In the other two quadrants, higher (or lower) cost levels have to be weighed against greater (or lesser) QALY gains.

The cost-effectiveness acceptability curve (Van Hout et al 1994; Barrett and Byford 2003) can be used as another way to visualize the cost-effectiveness results. Such an acceptability curve represents the probability that the intervention is cost-effective relative to the control condition, given varying willingness-to-pay ceilings for gaining one QALY (see **figure 2.3**).

Sensitivity analyses

Uncertainty was introduced in the trial data due to high dropout rates. Within the four conditions 33, 50, 43 and 61 participants reported costs and/or utility data at the 12-month follow up (t_3), leading to 69%, 54%, 60% and 45% dropout rates in the intervention with no support, support on request, weekly support and the control condition, respectively. Dropout rates after 6 weeks (t_1) and 3 months (t_2) were 35%, 25%, 30% and 12% and 49%, 37%, 51% and 28%, respectively. In this context, results could depend on the chosen imputation method. Therefore, we conducted sensitivity analyses not only using expectation maximization (for the main analysis), but also regression imputation and last observation carried forward (LOCF) to see how robust results were under the different imputation methods.

Next to that, we tested the robustness of our conclusions with respect to the method chosen to correct for baseline differences. As an alternative to the earlier described method, we corrected for baseline differences using baseline adjusted

multivariate regression analysis and baseline adjusted seemingly unrelated regression equation (SURE) analysis.

Results

Sample

Participants were aged 44 years on average, predominantly female (65%), with a paid job (71%), but few were in treatment (13%) or on medication (9%) at the start of the trial. Baseline utility was 0.68 on average, CES-D scores averaged 27.14, HADS-A scores averaged 10.13, whereas baseline total costs in the past four weeks were €704 on average. Total costs at baseline were similar in the intervention arms, but appeared to be (non-significantly) lower in the control condition (**table 2.3**). Twenty-three percent of the participants who were randomized to some form of coaching chose chat sessions, 73% chose email support, 2% switched between chat and email, and 4% dropped out of the treatment before they had made a decision.

Table 2.3. Descriptive statistics respondents at baseline

Variable	Overall	Condition 1	Condition 2	Condition 3	Condition 4
N	431	107	108	106	110
Female (%)	65%	64%	64%	68%	65%
Age (mean)	44.25	42.75	44.35	42.85	46.95
Income (per hour)	€ 28.97	€ 27.77	€ 29.39	€ 28.65	€ 30.03
In treatment (%)	13%	9%	13%	16%	15%
Medication (%)	9%	7%	9%	9%	12%
Paid job (%)	71%	76%	73%	71%	63%
Working hours per week	29.96	30.98	29.74	29.03	30.02
On sick leave (%)	23%	23%	22%	31%	15%
Total costs	€ 704	€ 736	€ 733	€ 722	€ 627
Utility	0.68	0.67	0.67	0.69	0.69
CES-D baseline	27.14	27.24	26.53	27.25	27.56
HADS-A baseline	10.13	10.37	9.82	10.00	10.30

Conditions: 1 = problem solving without coaching; 2 = problem solving with on-demand coaching; 3 = problem solving with weekly coaching; 4 = no problem solving and non-specific coaching.

Clinical effects on depression and anxiety

In the short term (3-months), CES-D and HADS-A outcomes were similar in the different conditions. CES-D equaled 19.9 (SD 8.0), 18.8 (SD 7.5), 18.9 (SD 8.4) and 19.3 (SD 9.7) in the intervention conditions with no support, support on request and weekly support and the control condition, respectively. HADS-A equaled 7.6 (SD 3.1), 7.3 (SD 3.0), 7.4 (SD 3.0) and 7.5 (SD 3.8) in the different conditions, respectively.

Twelve months after the start of the intervention, CES-D outcomes were 20.1 (SD 7.5), 18.4 (SD 7.3), 18.8 (SD 7.6) and 18.4 (SD 8.6) in the intervention conditions with no support, support on request and weekly support and the control condition, respectively. HADS-A outcomes were 7.4 (SD 3.0), 6.5 (SD 2.8), 7.2 (SD 3.1) and 6.8 (SD 3.4) in the different conditions, respectively. Interestingly, the intervention arms do not show better outcomes than the control condition (non-specific coaching). Although not significant, it is the control condition that generally shows better outcomes in both the CES-D (t=1.532, df=215, p=0.937; t=0.009, df=216, p=0.504; t=0.342, df=214, p=0.634, respectively) and HADS-A (t=1.499, df=215, p=0.933; t=-0.513, df=216, p=0.305; t=0.928, df=214, p=0.823, respectively). In terms of clinical outcomes over the 12-month period, the intervention cannot be seen as a favorable alternative to the control condition with non-specific support. It is important to note that this particular control condition with non-specific coaching is expected to generate more effects than what could be expected when comparing to a waitlist condition, thus diminishing the incremental effects as shown by the intervention conditions.

QALY gains

At 12 months, the mean per-participant cumulative QALYs were equal to 0.732 (SD 0.15) in the no support condition, 0.742 (SD 0.13) in the support on request condition, 0.726 (SD 0.14) in the weekly support condition and 0.723 (SD 0.13) in the non-specific chat or email condition. Incremental QALY gains were 0.009 (SD 0.023), 0.018 (SD 0.022) and 0.003 (SD 0.023) in the intervention conditions with no support, support on request and weekly support, respectively. As expected, one-sided t-tests show that these differences are non-significant (t=0.469, df=215, p=0.320; t=1.040, df=216, p=0.150; t=0.134, df=214, p=0.447, respectively). The probability that QALY gains are higher than in the control condition was 66%, 80% and 54% for the different intervention arms, respectively.

Costs

Health care utilization and corresponding health care costs are presented in **table 2.4** and **table 2.5**, which present the descriptive statistics of the observed data before imputation.

Table 2.4. Mean (standard error) health care visits as observed in the data (before imputation)

		Condition 1	Condition 2	Condition 3	Condition 4
General practitioner	T0	0.50 (0.09)	0.54 (0.10)	0.43 (0.08)	0.43 (0.08)
	T2	0.38 (0.10)	0.41 (0.14)	0.48 (0.20)	0.34 (0.08)
	T3	0.43 (0.18)	0.26 (0.08)	0.26 (0.11)	0.46 (0.12)
Company doctor	T0	0.07 (0.03)	0.07 (0.03)	0.07 (0.03)	0.07 (0.03)
	T2	0.07 (0.04)	0.06 (0.03)	0.08 (0.05)	0.09 (0.04)
	T3	0.03 (0.03)	0.04 (0.03)	0.11 (0.09)	0.05 (0.04)
Social worker	T0	0.09 (0.05)	0.04 (0.02)	0.04 (0.02)	0.05 (0.02)
	T2	0.02 (0.02)	0.03 (0.02)	0.04 (0.04)	0.05 (0.03)
	T3	0.00 (0.00)	0.00 (0.00)	0.06 (0.05)	0.02 (0.02)
Regional mental health service	T0	0.07 (0.02)	0.25 (0.1)	0.15 (0.06)	0.07 (0.03)
	T2	0.09 (0.05)	0.19 (0.1)	0.15 (0.06)	0.16 (0.06)
	T3	0.06 (0.04)	0.42 (0.23)	0.26 (0.11)	0.14 (0.07)
Alcohol and drug consultant (CAD)	T0	0.04 (0.04)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	T2	0.00 (0.00)	0.00 (0.00)	0.02 (0.02)	0.00 (0.00)
	T3	0.00 (0.00)	0.02 (0.02)	0.02 (0.02)	0.00 (0.00)
Independent psychotherapist	T0	0.06 (0.03)	0.14 (0.05)	0.14 (0.05)	0.08 (0.03)
	T2	0.20 (0.11)	0.15 (0.07)	0.06 (0.04)	0.31 (0.08)
	T3	0.29 (0.15)	0.26 (0.16)	0.28 (0.12)	0.19 (0.07)
Physiotherapist	T0	0.25 (0.08)	0.58 (0.18)	0.48 (0.16)	0.87 (0.25)
	T2	0.27 (0.14)	0.49 (0.20)	0.75 (0.30)	0.24 (0.09)
	T3	0.31 (0.19)	0.81 (0.29)	0.64 (0.24)	0.62 (0.21)
Medical specialist	T0	0.20 (0.05)	0.19 (0.05)	0.28 (0.07)	0.21 (0.05)
	T2	0.15 (0.08)	0.24 (0.08)	0.29 (0.12)	0.26 (0.08)
	T3	0.11 (0.05)	0.26 (0.10)	0.34 (0.14)	0.19 (0.07)
Complementary and Alternative Medicine	T0	0.36 (0.20)	0.36 (0.11)	0.40 (0.10)	0.40 (0.15)
	T2	0.24 (0.14)	0.26 (0.10)	0.15 (0.08)	0.28 (0.09)
	T3	0.17 (0.09)	0.19 (0.09)	0.11 (0.09)	0.56 (0.26)
Day treatment	T0	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
	T2	0.00 (0.00)	0.00 (0.00)	0.08 (0.08)	0.00 (0.00)
	T3	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

Table 2.5. Costs in euros (2011 prices) as observed in the data (before imputation)

		Condition 1	Condition 2	Condition 3	Condition 4
General practitioner	T0	€ 14.50	€ 15.72	€ 12.70	€ 12.58
	T2	€ 11.18	€ 12.05	€ 14.07	€ 9.88
	T3	€ 12.54	€ 7.73	€ 7.47	€ 13.47
Company doctor	T0	€ 4.78	€ 5.42	€ 4.83	€ 4.78
	T2	€ 5.32	€ 4.30	€ 5.62	€ 6.40
	T3	€ 2.09	€ 2.76	€ 7.78	€ 3.48
Social worker	T0	€ 6.35	€ 2.52	€ 2.56	€ 3.18
	T2	€ 1.24	€ 2.00	€ 2.61	€ 3.40
	T3	€ 0.00	€ 0.00	€ 4.34	€ 1.08
Regional mental health service	T0	€ 11.70	€ 44.70	€ 26.99	€ 11.70
	T2	€ 16.25	€ 34.18	€ 27.50	€ 29.05
	T3	€ 10.22	€ 74.21	€ 45.65	€ 25.54
Alcohol and drug consultant (CAD)	T0	€ 6.68	€ 0.00	€ 0.00	€ 0.00
	T2	€ 0.00	€ 0.00	€ 3.44	€ 0.00
	T3	€ 0.00	€ 3.37	€ 3.80	€ 0.00
Independent psychotherapist	T0	€ 4.51	€ 11.18	€ 11.39	€ 6.77
	T2	€ 16.10	€ 11.84	€ 4.64	€ 25.16
	T3	€ 23.00	€ 21.26	€ 22.27	€ 15.33
Physiotherapist	T0	€ 9.50	€ 21.96	€ 18.11	€ 32.72
	T2	€ 10.27	€ 18.27	€ 28.23	€ 8.94
	T3	€ 11.83	€ 30.54	€ 24.03	€ 23.30
Medical specialist general hospital	T0	€ 14.77	€ 13.94	€ 21.31	€ 15.48
	T2	€ 10.95	€ 17.71	€ 21.72	€ 19.76
	T3	€ 8.60	€ 19.89	€ 25.63	€ 14.34
Complementary and Alternative Medicine	T0	€ 19.32	€ 19.14	€ 21.00	€ 21.30
	T2	€ 12.53	€ 14.03	€ 8.15	€ 14.58
	T3	€ 9.09	€ 10.00	€ 5.64	€ 29.44
Day treatment	T0	€ 0.00	€ 0.00	€ 0.00	€ 0.00
	T2	€ 0.00	€ 0.00	€ 12.38	€ 0.00
	T3	€ 0.00	€ 0.00	€ 0.00	€ 0.00
Out of pocket costs	T0	€ 6.23	€ 8.23	€ 7.90	€ 8.15
	T2	€ 5.55	€ 7.22	€ 8.15	€ 7.12
	T3	€ 5.40	€ 8.56	€ 7.60	€ 8.39
Productivity costs	T0	€ 637.40	€ 591.81	€ 595.07	€ 503.37
	T2	€ 350.64	€ 668.22	€ 428.11	€ 329.89
	T3	€ 294.69	€ 386.74	€ 418.10	€ 249.53
Pharmaceuticals	T0	€ 1.02	€ 1.02	€ 1.38	€ 2.10
	T2	€ 1.72	€ 2.16	€ 2.74	€ 4.76
	T3	€ 4.36	€ 19.80	€ 5.69	€ 26.12
Total costs	T0	€ 735.70	€ 733.48	€ 722.27	€ 626.95
	T2	€ 408.81	€ 792.85	€ 554.80	€ 456.28
	T3	€ 392.49	€ 588.67	€ 397.94	€ 414.35
Intervention costs		€ 56.88	€ 155.09	€ 324.14	€0.00

Costs were imputed and corrected for baseline differences and linear interpolation was used to arrive at yearly cost estimates. This resulted in mean per-participant cumulative costs over one year of €6,081 (SD 851), €7,542 (SD 962), €6,848 (SD 943) and €7,072 (SD 949) in the no support, support on request, weekly support and control group, respectively. Incremental costs relative to the control condition are then -€991 (SD 1,266), €469 (SD 1,352) and -€225 (SD 1,341) for the intervention conditions with no support, support on request and weekly support, respectively. The probability that per-participant costs are more favourable than in the control group was 79%, 35% and 56% in the intervention arms, respectively.

Cost-utility

The ICER for the intervention condition with support on request was equal to €469 / 0.018 = €25,697 per QALY gained. The intervention conditions with no support and weekly support dominated the control condition with average cost savings of €991 and €225, and average QALY gains of 0.009 and 0.003, respectively.

The incremental cost-effectiveness ratios are surrounded by uncertainty. **Figure 2.2** presents the cost-effectiveness plane for the intervention conditions versus the control condition. The incremental costs are plotted on the y-axis and the incremental effects on the x-axis. Each dot (n=2,500) represents a bootstrap replication of the incremental cost-effectiveness ratio. **Table 2.6** presents the distribution of the bootstrap replications over the four different quadrants in the ICER plane for each of the intervention arms. For the intervention with no support, 56% of the dots are in the lower right-hand quadrant; 23% in the lower left-hand quadrant, 9% in the upper right-hand quadrant, and 11% in the upper left-hand quadrant, indicating a 79% probability that the intervention is less expensive, and a 66% probability that the intervention is more effective than the control condition.
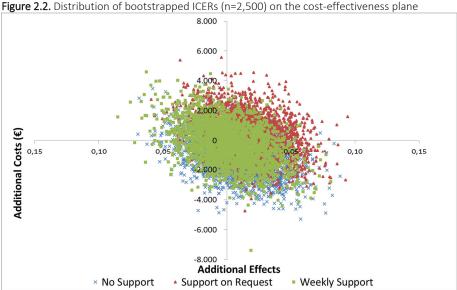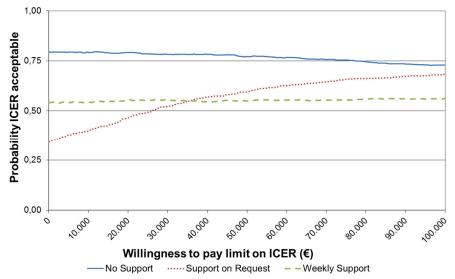
Figure 2.2. Distribution of bootstrapped ICERs (n=2,500) on the cost-effectiveness plane

Table 2.6. Bootstrap outcomes: incremental cost-effectiveness for different intervention arms

	No Support	Support on Request	Weekly Support
Incremental costs, €	-991	469	-225
Incremental QALY gains	0.009	0.018	0.003
ICER, € (median)	-33,798	1,480	-27,877
Distribution on the cost-effectiveness plane			
. 1st quadrant (north-east)	0.09	0.48	0.16
. 2nd quadrant (inferior: north-west)	0.11	0.17	0.27
. 3rd quadrant (south-west)	0.23	0.03	0.19
. 4th quadrant (dominant: south-east)	0.56	0.32	0.37
Percentage of ICERs below WTP ceiling			
. €0	79	34	54
. €10,000	79	40	54
. €20,000	79	47	56
. €30,000	78	52	55
Notes: (ICER) incremental cost-effectiveness ratio; (WTP) willingness-to-pay.			

Acceptability

Figure 2.3 presents the cost-effectiveness acceptability curve for each of the intervention conditions versus the control condition. The intervention condition with no support has a high probability of being more cost-effective than the control condition over the considered range of willingness-to-pay threshold values per QALY. The intervention condition with support on request exceeds the 50% probability of being more cost-effective than the control condition at the WTP threshold value of

€26,100. The intervention condition with weekly support is slightly but consistently above the 50% probability of being more cost-effective than the control condition.

Figure 2.3. ICER acceptability curve: probability that the intervention is more cost-effective than the control condition (y-axis) given varying thresholds for willingness-to-pay (x-axis), based on 2,500 bootstrap replications

Sensitivity analyses

When imputing missing data using different imputation methods, some conclusions may change. However, for each imputation method, the intervention condition with no support is preferred over the considered range of WTP threshold values when compared to the control condition. When imputing missing values using LOCF, the condition with support on request is preferred over the control condition for the entire range of WTP values, whereas the condition with weekly support does not exceed the 50% probability of being more cost-effective than the control condition until a WTP value of €50,700. When imputing missing values using regression, the condition with support on request exceeds the 50% probability of being more cost-effective than the control condition at WTP values starting from €32,700, whereas the intervention condition with weekly support has a 50% or higher probability of being more cost-effective than the control condition for nearly the full range of WTP values considered (up to WTP values of €89,900). Overall, our conclusions regarding the intervention condition with no support are robust, whereas conclusions regarding the intervention conditions with support on request and weekly support depend on the imputation method.

Our conclusions are robust under different methods of baseline correction, such as baseline adjusted multivariate regression analysis and baseline adjusted seemingly unrelated regression equation (SURE) analysis.

DISCUSSION

Main findings

Cumulative QALY outcomes in the three intervention arms had a 66%, 80% and 54% probability of being higher than in the control condition. The three intervention arms had a 79%, 35% and 56% probability of having lower cumulative costs than the control condition.

The intervention condition without support was generally associated with a high probability of being more cost-effective than the control condition. The intervention conditions with weekly support and support on request were not consistently associated with a probability higher than 50% of being more cost-effective than the control condition for the different imputation methods used. In terms of clinical effect, the intervention conditions did not show better outcomes than the control condition over the 12-month period.

Limitations

There are several limitations to this study. An important limitation is the high study dropout rate of 45%-69% in the 12-month follow-up for the different conditions. High dropout rates lead to a greater potential bias in the estimates of average change in utility and costs across the different arms in the trial, which can impact on the conclusions. The impact of high dropout rates was minimized using various imputation techniques and evaluating the robustness of the conclusions in sensitivity analyses. Apart from the intervention with no support, results depend on the imputation method chosen. Due to overall high dropout rates, results are therefore best regarded as tentative. The adherence to the Internet-based interventions was low, as is common with Internet-based interventions (Christensen et al 2009), which means that most participants did not receive a high dose of the intervention and this may have influenced the results.

Next, it was not possible to conduct a double-blinded trial. This is true for most randomized trials of psychological interventions, but it may nevertheless have distorted the outcomes of our trial.

Another limitation is the lack of an additional waitlist control condition. No resource use data were collected in the waitlist control condition at any of the

measurements and no clinical data were collected after the post-test. The use of a non-specific coaching control condition is likely to lead to smaller incremental health effects in the intervention arms. Indeed, when analyzing clinical effects at post-test (t_1), we find significant effects in the intervention conditions compared to the waitlist condition, and non-significant effects when comparing the intervention conditions with the non-specific coaching condition. When an intervention is not associated with clinical effects, it is common to focus solely on costs by performing a cost-minimization analysis rather than a cost-utility analysis. This approach should be taken when evidence regarding the clinical equivalence is established through a properly powered non-inferiority trial (as opposed to our trial, which was designed as a superiority trial). Although the lack of evidence regarding clinical effects could be seen as an argument for performing a cost-minimization analysis, our superiority design, as well as the clinical effects found when comparing the intervention conditions with the waitlist condition at post-test, strengthened the case for performing a cost-utility analysis instead.

In this cost-utility study direct medical costs and productivity losses were considered, where productivity costs made up for most of the total costs. Preferably, the societal perspective considers a more diverse range of cost types, for example costs regarding household, family or caregivers, in order to provide a more complete picture of the cross-sectorial impact of the intervention on society. Our study, however, was limited by the available types of costing data.

Resource use was collected at baseline and after three and 12 months, and then used to estimate cumulative yearly costs by correcting for baseline differences, linear interpolation and the AUC method. Resource use was furthermore identified by a retrospective interview, which could have introduced recall bias. Ideally, resource use would have been collected at smaller, more regular intervals, in order to minimize uncertainty around costs and changes therein.

The trial was conducted in the Netherlands. Although it is expected that results would apply to different contexts as well, further research is required to confirm this.

It is unclear how the intervention impacts on costs and effects beyond the 12 months of the trial. Longer-term follow-ups or a modelling approach are needed to estimate the impact of the intervention beyond 12 months.

Finally, we used questionnaires to assess symptoms of anxiety and depression and did not perform diagnostic interviews to diagnose anxiety and depression. Therefore we do not know to what extent we were able to prevent the onset of depression and anxiety.

The wider context

Depression and anxiety are highly disabling conditions both in terms of population health and in terms of economic impact. The large gap between the disease burden in the population and the disease burden averted by treatment (Andrews et al 2004) highlights the importance of developing new ways to treat people suffering from (symptoms of) depression and anxiety. Current economic as well as demographic developments make it necessary to look for cost-effective interventions that put minimal pressure on financial and labour resources. Although our results at this stage are only tentative, it is encouraging to see that the intervention conditions with higher involvement of coaches are not necessarily the conditions with the most favourable cost-effectiveness ratios. Intervention formats involving less time from health care professionals may increase the number of patients treated when human resource capacity is constrained.

Implications

For health care professionals, our results may implicate that their therapeutic approach might have to be geared towards providing blended therapy more often. After all, providing more e-health could potentially increase the number of patients treated under an equal budget, without compromising the effectiveness of the therapy. However, before such a conclusion can be made, further research is needed to confirm that our results hold across various groups of patients, and to better guarantee the patients' adherence with online interventions.

For policymakers, our research lends support to the idea that more patients can receive treatment under equal budget. Implementation of such a change, however, requires the support of the stakeholders involved. When convinced of the desirability of this type of intervention, policymakers need to translate the research findings to the local context and take stakeholder's perspectives into account to increase the likelihood of successful implementation.

Chapter 3

Modelling the cost-effectiveness of health care systems for alcohol use disorders: how implementation of eHealth interventions improves cost-effectiveness

Based on our publication in the Journal of Medical Internet Research

Abstract

Background: Informing policy decisions about the cost-effectiveness of health care systems (i.e., packages of clinical interventions) is probably best done using a modelling approach. To this end, an alcohol model (ALCMOD) was developed.

Objective: The aim of ALCMOD is to estimate the cost-effectiveness of competing health care systems in curbing alcohol use at the national level. This is illustrated for scenarios where new eHealth technologies for alcohol use disorders are introduced in the Dutch health care system.

Method: ALCMOD assesses short-term (12-month) incremental cost-effectiveness in terms of reductions in disease burden, that is, disability adjusted life years (DALYs) and health care budget impacts.

Results: Introduction of new eHealth technologies would substantially increase the cost-effectiveness of the Dutch health care system for alcohol use disorders: every euro spent under the current system returns a value of about the same size (€1.08, i.e., a "surplus" of 8 euro cents) while the new health care system offers much better returns on investment, that is, every euro spent generates €1.62 in health-related value.

Conclusion: Based on the best available evidence, ALCMOD's computations suggest that implementation of new eHealth technologies would make the Dutch health care system more cost-effective. This type of information may help (1) to identify opportunities for system innovation, (2) to set agendas for further research, and (3) to inform policy decisions about resource allocation.

Introduction

Alcohol use disorders are a leading cause of disease burden (Ezzati et al 2004; Rehm et al 2006) and are associated with substantial economic costs (Smit et al 2006a; Rehm et al 2009; Mohapatra et al 2010). Therefore, curbing alcohol use has long been recognized as an important public health objective (World Health Organization 2008; Beaglehole and Bonita 2009). Health care systems play a crucial role in achieving this objective, but most health care systems offer room for improvement in terms of greater efficiency. This begs the question what type of health care system (i.e., what mix of interventions) is optimal. We could provisionally define an optimal health care system in terms of meeting the following criteria: the health care system needs to be acceptable to its recipients, appropriate to its practitioners, scalable to absorb increasing demands for health care, effective to generate the required health gains, and economically affordable to become sustainable over time. Public health planners need ways to design health care systems that optimize these criteria, compare the relative advantage of newly designed systems with the current one, and choose the most cost-effective system. This is a daunting but important task. However, this task might be facilitated with a simulation model, which can compare a "base case" scenario (e.g., the current mix of clinical interventions) with an alternative (hypothetical) scenario consisting of new interventions or a different mix of interventions. In order for it to be helpful, the model should be able to evaluate the relative advantage of one system over another in terms of incremental cost-effectiveness and be used as an aid to decision-making. With these aims in mind, we developed an alcohol model (ALCMOD) that can address the above issues. Developing ALCMOD was conducted within the framework of the World Health Organization's International Action Plan on Implementing eHealth Technologies for Substance Abuse. In this context, we wanted to shed light on the population-level cost-effectiveness of health care systems for alcohol use disorders before and after the introduction of new eHealth technologies in Belarus, Brazil, India, Mexico, and the Netherlands. ALCMOD is programmed in Microsoft Excel 2007, because Excel is available on most computers.

The purpose of this paper is to describe ALCMOD's input and output and to take an in-depth look at the model's throughput: its computational strategies, the underlying assumptions, and its limitations. One such limitation is ALCMOD's focus on short-term impacts. Restricting the time horizon to 1 year was a conscious choice because there are several alcohol use disorders (heavy, hazardous, and harmful use and

alcohol dependence; see **textbox 3.1** for definitions) and a lack of empirical data that help to quantify the longer-term treatment effects and relapse rates for each of the alcohol use disorders. By contrast, simulation of short-term health and budget impacts is straightforward and requires fewer assumptions. Strengths of ALCMOD include its ability to evaluate combinations of interventions, its adaptability to different populations and settings, its capacity to handle uncertainty in input parameters, and the way it incorporates coverage and adherence rates for each of the modelled interventions. We illustrate ALCMOD's computations for the base case of usual care in the Netherlands versus an alternative scenario consisting of usual care augmented with three eHealth interventions: the DrinkTest, DrinkingLess, and an online therapist-led treatment for problem drinking, termed OnlineTreatment henceforth. The DrinkTest is a brief online intervention consisting of screening one's alcohol use followed by automated personalized advice. DrinkingLess is an online four-step cognitive behavioural intervention. The steps in DrinkingLess are: (1) exploring one's alcohol use, (2) goal setting, (3) behavioural change, and (4) maintenance. Both the DrinkTest and DrinkingLess have been evaluated in randomized trials and meta-analytically and were found to be effective in curbing alcohol use (Meijer et al 2006; Riper et al 2008; 2011; Boon et al 2011). Both the DrinkTest and DrinkingLess are pure self-help interventions, but OnlineTreatment is a therapist-led intervention. Communication between participant and therapist is conducted over the Internet in seven synchronous written chat sessions of 45 minutes each. The sessions are thematically structured and cover themes such as goal setting, self-control techniques, monitoring, recognizing situations that incur a risk of relapse, and relapse prevention techniques.

Textbox 3.1. Description alcohol use disorders

Alcohol use disorders from the lexicon of alcohol and drug terms published by the World Health Organization (1994):
- Abstinence is defined as refraining from drinking alcoholic beverages.
- Moderate drinking is the consumption of alcohol that does not exceed guidelines for moderate drinking in terms of volume or quantity per occasion.
- Heavy drinking is defined as drinking in excess of the standard of moderate drinking (see moderate drinking, above).
- Hazardous use (*International Classification of Disease, Tenth Revision* [*ICD-10*] code Z72.1) is a pattern of heavy drinking and/or binge drinking that carries with it a risk of harmful consequences to the drinker. These consequences may be detrimental to physical or mental health or have adverse social consequences to the drinker or others. Other potential consequences include worsening of existing medical conditions or psychiatric illnesses, injuries caused to self or others due to impaired judgment after drinking, high-risk sexual behaviours while intoxicated, and worsening of personal or social interactions.
- Harmful drinking (*ICD-10* code F10.1) is a pattern of drinking that is causing damage to health. The damage may be either physical (e.g., liver cirrhosis from chronic drinking) or mental (e.g., depressive episodes secondary to drinking). Harmful patterns of use are often criticized by others and are sometimes associated with adverse social consequences of various kinds. Harmful drinking has persisted for at least 1 month or has occurred repeatedly over the past 12-month period; subject does not meet criteria for alcohol dependence.
- Alcohol dependence (*ICD-10* code F10.2) At least 3 of the following criteria are met: tolerance; withdrawal symptoms; impaired control; preoccupation with acquisition and/or use; persistent desire or unsuccessful efforts to quit; sustains social, occupational, or recreational disability; and use continues despite adverse consequences.

OnlineTreatment has been evaluated in a randomized trial (Blankers et al 2011). Preliminary results (Blankers et al 2012) indicate that OnlineTreatment is effective and cost-effective. It is worth noting that the three eHealth interventions increase in intensity and could be used in a stepped-care framework, thus starting with the least intensive intervention, the DrinkTest, and moving up to the more intensive levels of DrinkingLess and OnlineTreatment, if so required. The emergence of evidence-based eHealth technologies offers opportunities for innovation in existing health care systems. The new technologies may help to reach population segments that were hitherto not reached because they live in hard to reach rural areas or because they may have shied away from face-to-face delivered health services out of fear of stigma. The new technologies are also very scalable, thus allowing people to access

health care services in an unprecedented way. In addition, the new eHealth technologies could be cost-effective, especially when offered as well-structured self-help interventions or as interventions with (minimal) therapist support. Considering the global health gap with regard to the alcohol use disorders (World Health Organization 2008), these developments could become quite important. However, to date there is only limited evidence for the cost-effectiveness of eHealth interventions (Tate et al 2009). For these reasons, it is opportune to conduct a population-level health economic evaluation of the possible health gains and budget impacts of adding new eHealth technologies to the existing health care system for alcohol use disorders.

Methods

Target population
By way of input, ALCMOD requires data that describe key characteristics of the target population. Selecting the name of the country will automatically trigger ALCMOD to upload the age and gender distribution of the population of the selected country and the corresponding mortality rates. ALCMOD also needs to know the size of the target population, and in the Netherlands, the target population consists of 993,200 men and 222,800 women aged 18 to 69 years who could be classified as *problem drinkers* (estimates based on a sample of approximately 7,500 respondents (van Dijck and Knibbe 2005)). Other required input is the pre-intervention profile of the target population based on the Alcohol Use Disorders Identification Test, the AUDIT (Higgins et al 1976; Reid et al 1999; Babor et al 2001; Room et al 2005). The decision to base ALCMOD on the AUDIT was motivated by the idea that the AUDIT is globally used. Moreover, newly developed eHealth interventions commissioned by WHO are most likely to screen participants with the AUDIT. Thus, even when a country has no AUDIT data yet, these data are likely to become available via eHealth interventions in the near future. In the Netherlands, data from the AUDIT are available and can be automatically uploaded in ALCMOD.

Intervention packages
ALCMOD allows a description of the intervention mix representing the base case scenario and the designing of an alternative scenario with a different mix of interventions or new interventions added to existing ones. In ALCMOD's default setting, a range of interventions — both face-to-face interventions and eHealth

interventions — are shown for heavy, hazardous, and harmful alcohol use and alcohol dependence. Two parameters need to be set for each of the interventions: the coverage rate and the adherence rate.

Coverage rate

When some of the interventions shown in ALCMOD's default setting are not available in a country, then their coverage rate has to be set to 0%. This is equivalent to saying that the intervention is not offered in a country. Other interventions might be available for every person belonging to the target population (the population eligible to receive the intervention), and the coverage rate is then set to 100% (universal coverage). However, due to the many obstacles to full implementation, the coverage rate of most interventions is somewhere between 0 and 100% and can be set accordingly in ALCMOD.

Adherence rate

Recipients of interventions might be less than willing or unable to fully comply with the intervention, and the degree of adherence is likely to moderate treatment response. Therefore, the adherence rate is an important parameter when evaluating the effectiveness of interventions. Adherence rates may be obtained from the literature, experts, or via focus groups in the target population.

The idea is that health care scenarios can be developed by changing the level of coverage for a series of interventions. **Table 3.1** shows the settings for the three scenarios that we modelled: (1) the current Dutch health care system for alcohol use disorders without eHealth interventions (base case scenario), (2) the Dutch health care system augmented with the eHealth interventions (alternative scenario 1), and (3) the Dutch health care system where face-to-face interventions have been substituted for 50% by the new eHealth interventions (alternative scenario 2).

Table 3.1: Modelled scenarios: coverage rates (%) for each of the interventions

Target group Alcohol use disorder	Intervention	Base case scenario	Alternative scenario 1	Alternative scenario 2
Heavy	Brief face-to-face intervention [a]	10	10	5
	Online brief intervention [b]	0	5	5
Hazardous	Brief face-to-face intervention [a]	10	10	5
	Online brief intervention [b]	0	5	5
	Behavioural intervention [c]	6	6	3
	Online behavioural intervention [d]	0	6	3
Harmful	Behavioural intervention [c]	9	9	3
	Online behavioural intervention [d]	0	9	3
	Online therapist-led treatment [e]	0	9	3
	Detox and acamprosate [f]	5	5	5
	Aftercare and rehab with AA [g]	5	5	5
Dependence	Behavioural intervention [c]	5	5	2.5
	Online therapist-led treatment [e]	0	5	2.5
	Detox and acamprosate [f]	5	5	5
	Aftercare and rehab with AA [g]	5	5	5

[a] Brief face-to-face is modelled as a brief intervention consisting of screening followed by personalised feedback by a physician usually in a single session (<10 min), occasionally in two sessions (one for screening, the other for personalised feedback).
[b] Online brief intervention is modelled as online screening and automated personalised feedback (DrinkTest).
[c] Behavioural intervention is modelled as eight to ten sessions of individual cognitive behavioural therapy (CBT) under the guidance of a therapist, followed by one booster session.
[d] Online self-help intervention (DrinkingLess) is modelled as four (range 3 to 12) sessions of online interactive CBT-based self-help preceded by referral by a general practitioner (GP).
[e] Online therapist-led intervention is modelled as eight sessions of online therapist-led CBT.
[f] Detox is modelled as one week ambulatory detoxification followed by clinical management with acamprosate.
[g] Aftercare and rehabilitation is modelled as participation in Alcoholics Anonymous (AA) over 12 months.

The choice of intervention mix was informed by Room et al. (2005) and Benegal et al. (2009) and the Dutch multidisciplinary guideline for the treatment of alcohol use disorders (Landelijk Kwaliteitsinstituut CBO and Trimbos Instituut 2009). The choice of interventions was also motivated by two additional considerations: availability of evidence of the intervention's effectiveness in the meta-analytical literature (Moyer et al 2002) and the non-overlapping *independent* nature of the interventions such that each intervention could be added to other interventions without creating overlap for a specific alcohol use disorder.

Finally, the scenarios have been simplified by assuming that all interventions are associated with an adherence rate of 50%. This was done to ensure that differences in the cost-effectiveness ratios are due to fundamental differences in health technologies, ruling out the effect of greater or lesser treatment adherence.

However, it is possible to adjust adherence rates in ALCMOD. After all, some interventions might be associated with better or poorer adherence, and adherence itself might be amenable to intervention such as motivational enhancement. Changing the adherence parameters allows evaluation of these issues.

Cost and effect parameters
In the ALCMOD default settings, some of the intervention parameters have been preset and need not be changed, but can be changed if so required. These parameters are the costs and the effects of the interventions.

Costs
ALCMOD's default setting makes use of the full economic cost price of each of the interventions. To be precise, the costs are the per-participant costs of delivering an intervention expressed in euro (€) for the Netherlands in the reference year 2009 (see **table 3.2**). The costs are based on the amount of resources (labour, facilities, and supplies) used for offering the intervention during its postimplementation stage. We made our own costing tool to estimate the costs (in euro) of interventions in a systematic and uniform way that is in agreement with the Dutch guideline for costing health care interventions (Oostenbrink et al 2004). For other countries, the per-participant costs of offering an intervention need to be assessed. These assessments can be carried out with the help of an auxiliary costing tool, for example Cost It, available from WHO's CHOICE website. Neither costs nor effects are discounted because ALCMOD takes a short-term (12-month) perspective.

Table 3.2: Per-patient intervention costs in 2009 euro (€) with uncertainty range

Target group Alcohol use disorder	Intervention	Costs, €	Uncertainty range* Low	High
Heavy	Brief face-to-face intervention [a]	58	52	75
	Online brief intervention [b]	10	9	10
Hazardous	Brief face-to-face intervention [a]	58	52	75
	Online brief intervention [b]	10	9	10
	Behavioural intervention [c]	2,024	1,702	2,550
	Online self-help intervention [d]	207	198	224
Harmful	Behavioural intervention [c]	2,024	1,702	2,550
	Online self-help intervention [d]	207	198	224
	Online therapist-led intervention [e]	764	227	1,451
	Detox and acamprosate [f]	1,800	1,620	2,232
	Aftercare and rehab with AA [g]	500	250	750
Dependence	Behavioural intervention [c]	2,024	1,702	2,550
	Online therapist-led intervention [e]	1,276	979	1,408
	Detox and acamprosate [f]	1,800	1,620	2,232
	Aftercare and rehab with AA [g]	500	250	750

* Uncertainty range based on 1,000 simulations, assuming a gamma distribution.
[a] Brief face-to-face intervention modelled as screening at € 5.70 followed by 1 or 2 (Poisson distributed) 10-minute contacts with GP at €32.03 per contact.
[b] Online brief intervention (DrinkTest) modelled as 40% of target population (N=1,255,000), reached with information about the website, 8% responding to AUDIT screener and receiving automated personalised feedback. Per-participant annual costs include website upgrading at €50,000, research at €50,000 and hosting at €25,000.
[c] Behavioural intervention is modelled as 8 to 14 (Poisson distributed) sessions of cognitive behavioural therapy (CBT) under guidance by a therapist, including referral, intake and one booster session.
[d] Online self-help intervention (DrinkingLess) is modelled as 15% of target population (N= 1,255,000) reached with information about the website, 5% uptake rate, and 4 sessions (range 3 to 12) of online CBT-based self-help preceded by referral by a GP. Per-participant annual costs include €75,000 for website upgrading, €50,000 for research, €25,000 for hosting, plus €75,000 for forum moderation and technical assistance.
[e] Online therapist-led intervention is modelled as an average of 4 sessions (range 1 to 9) of 45 minutes each of online therapist-led CBT, preceded by GP referral. Per-participant costs include per annum costs of €8,000 for website upgrading, €5,000 for hosting, plus €2,000 for technical assistance.
[f] Detox is modelled as one week ambulatory detoxification followed by clinical management with acamprosate under the supervision of a substance use disorder treatment specialist and a physician over three months.
[g] Aftercare and rehabilitation is modelled as participation at Alcoholics Anonymous at an average of €500 (range €250 to €750) per patient for a year.

Effects

Intervention effects are expressed as the standardized mean difference, also known as Cohen's d. This metric indicates how many standard units (on a scale of standard deviations) the experimental group has improved relative to a control group on a relevant outcome such a change in drinking behaviour. The effect size d is often reported in the meta-analytical literature and gives access to a large body of scientific

evidence. We extracted effect sizes at 6- or 12-months follow-up for all the interventions from the meta-analytical literature and our own research (see **table 3.3**) and these values were used to populate ALCMOD with its default parameter settings.

Table 3.3: Effectiveness of the interventions: standardised mean differences (d), 95% confidence interval for d (95% CI) and references

Target group Alcohol use disorder	Intervention	d	95% CI
Heavy	Brief face-to-face intervention [a]	0.26	0.20 to 0.32
	Online brief intervention [b]	0.19	-0.02 to 0.40
Hazardous	Brief face-to-face intervention [c]	0.32	0.23 to 0.42
	Online brief intervention [b]	0.19	-0.02 to 0.40
	Behavioural intervention [d]	0.34	0.12 to 0.56
	Online self-help intervention [e]	0.31	-0.69 to 1.30
Harmful	Behavioural intervention [d]	0.34	0.12 to 0.56
	Online self-help intervention [e]	0.31	-0.69 to 1.30
	Online therapist-led intervention [f]	0.58	0.29 to 0.88
	Detox and acamprosate [g]	0.21	0.14 to 0.29
	Aftercare and rehab with AA [h]	0.28	0.20 to 0.37
Dependence	Behavioural intervention [i]	0.32	0.05 to 0.59
	Online therapist-led intervention [f]	0.59	0.30 to 0.90
	Detox and acamprosate [g]	0.21	0.14 to 0.29
	Aftercare and rehab with AA [h]	0.28	0.20 to 0.37

[a] *Moyer et al.'s (2002) meta-analysis of brief face-to-face interventions in approx. 4,300 users meeting criteria of at least heavy drinking.*
[b] *Randomized trial of 450 participants presenting with either excessive alcohol consumption (> 20 units weekly) and/or binge drinking (> 5 units on a single occasion on least one day per week) in the past 6 months (Meijer et al 2006).*
[c] *Reanalysis of Beich et al.'s (2003) meta-analysis of brief face- to-face interventions in 2,989 users meeting criteria of hazardous drinking.*
[d] *Walters' (2000) meta-analysis based on approximately 320 harmful users.*
[e] *Randomized trial of 261 excessive drinkers from the general population (Riper et al 2008) where odds ratio (OR) converted into d using Chinn's equation (Chinn 2000).*
[f] *Randomized trial of 250 adults with mean AUDIT score of 20 at baseline. Intervention was online treatment versus waitlist at 3 months with the AUDIT as outcome (Blankers et al 2011).*
[g] *Mann's (2004) meta-analysis of 1,670 people receiving acamprosate after detoxification where odds ratios converted into d using Chinn's method (Chinn 2000).*
[h] *Tonigan et al.'s (1996) meta-analysis of 2,097 harmful and dependent users where effect size r converted into d.*
[i] *Walters' (2000) meta-analysis based on approximately 210 dependent users.*

It is worth noting that ALCMOD uses two types of effects: the standardized mean difference, d, which was just discussed, and the impact of an intervention in terms of the percent reduction of pure alcohol intake in grams per day (g/day). The former effect (d) impacts on health-related quality of life (QoL) via changes in disorder severity. ALCMOD uses the percent reduction of pure alcohol intake to model treatment effects on mortality (see below for details). Although ALCMOD can handle different alcohol reduction rates for each of the modelled interventions, we have assumed a pre-post reduction of alcohol intake by 20% for all interventions (Moyer et al 2002; Davis et al 2002), because reduction of alcohol intake was not always reported in the literature. This should not overly distort outcomes because the short-term effects of alcohol use on mortality are small, thus limiting their impact on disease burden as measured by disability-adjusted life years (DALYs). We say this on the understanding that alcohol-related mortality becomes an important, even a dominant, factor when disease burden is modelled out to full life expectancy, especially in the more severe alcohol use disorders.

Here we need to address a final point about the required input for ALCMOD. ALCMOD can be operated in two modes: deterministic and probabilistic. In deterministic mode, ALCMOD does not take into account the uncertainty in parameters such as costs and effects. ALCMOD conducts all computations, but only once, and these calculations are primed on the mean value of all parameters. Much of ALCMOD's output, which is based on uncertainty, is then disabled. However, in probabilistic mode, ALCMOD can handle uncertainty surrounding the cost (in euro) and effect (d) parameters. Our costing tool assesses the uncertainty in costs with the help of simulations of resource use (with 1,000 iterations), and both randomized trials and meta-analyses of trials often report 95% confidence intervals of the effect size d. Thus we assume that costs are surrounded by an uncertainty range, and effects, by a 95% confidence interval, both having a lower and an upper limit. ALCMOD assumes a gamma distribution for costs and a normal distribution for the effect size d (in line with Briggs et al. (2006)). Both distributions can be specified in ALCMOD such that the distributions fit within the lower and upper limits of costs and effects. In probabilistic mode, ALCMOD then proceeds with drawing random values from these distributions, conducts all the computations, and repeats this process many times (maximum 10,000 times). This helps to capture uncertainty in the input parameters.

ALCMOD's Throughput

Differences in costs

Modelling cost differences between two health care systems is straightforward once the per-participant costs of delivering all modelled interventions have been estimated and when the coverage rates of the interventions have been established. The number of people in the target group (stratified by alcohol use disorder) is then multiplied by the coverage rate of each intervention and multiplied by the appropriate per-participant full economic cost price. The cost analyses are always conducted for both the base case and the alternative scenario, such that the cost difference between two modelled health care systems can be computed and expressed as incremental costs.

Differences in disease burden

The disability adjusted life year (DALY) is a measure of disease burden in a population. It combines two components of disease burden: morbidity and mortality. The first is related to lesser quality of life due to disability. Mortality arises when illness is associated with premature death. Thus, a DALY can be computed as the sum of years lost due to disability (YLD) plus years of life lost (YLL) due to mortality, hence, DALY = YLD + YLL.

The first term in the DALY equation, YLD, can be computed as the number of cases manifesting with an alcohol use disorder, N (point prevalence), weighted by a disability weight, DW. Thus, YLD = N × DW. DWs range from 0 to 1, where 0 is *no burden* (good health) and 1 refers to a health condition as undesirable as death. Although the literature offers advice for choosing DWs for the alcohol use disorders (Stouthard et al 1997; Murray et al 2000; Johns et al 2003; Kraemer et al 2005; Smit et al 2008), ALCMOD makes no use of DWs that are directly associated with each of the disorder-specific health states. Instead, it computes the (downward) shift in DW as a consequence of the treatment effect d. As said, d is the standardized mean difference indicating how many standard units the treatment group has moved away from the group that received no care. Thus, d is essentially a "health improvement shift" due to intervention. The task at hand, then, is to "translate" the health improvement shift (of size d) into a corresponding shift in DW. This strategy has been developed by Sanderson et al (2004) who used a panel of experts for obtaining a conversion factor of 0.18 (95% CI 0.16-0.20) to translate a shift in d into a shift in DW in alcohol use disorders. The change in DW is then multiplied by the appropriate number of people to arrive at an estimate of the number of YLD avoided.

When running in probabilistic mode, ALCMOD automatically conducts extensive uncertainty analyses around Sanderson's conversion factor. The second term in the DALY equation, YLL, is calculated as the difference in life expectancy when people reduce drinking levels. We obtained estimates of the gender-specific relative risk, RR, of all-cause mortality attributable to pure ethanol intake (in g/day) using the expression (Gmel et al 2003), $\ln(RR) = b_1*(\ln(x+1)) + b_2*\ln(x) + e$, where x = grams of pure ethanol intake per day and b_1 and b_2 are -0.1030 and 0.0035 for men and -0.0645 and 0.0029 for women, respectively (our own estimates from Gmel's paper (2003)). Exponentiating $\ln(RR)$ gives the relative risk, RR, and the RRs are then combined with the gender and age-specific mortality rates of the country for which the outcomes are modelled. This produces estimates of changes in life expectancy due to changes in alcohol intake. Because ALCMOD takes a short-term perspective, treatment induced impacts on life expectancy were calculated as the number of avoided deaths in the present year. The difference in YLD and YLL between the scenarios determines the difference in the disease burden as measured by DALY between two modelled health systems, the so-called incremental effects. ALCMOD offers the use of a (downward) *attenuation factor* that reduces the carry-over effects from lesser drinking to lesser mortality and better health-related quality of life. After all, it can be assumed that former drinkers still have a higher risk of dying and poorer quality of life than people who never drank before or have been consistent moderate drinkers (Barbosa et al 2010). In other words, returning to less risky drinking levels is assumed to be beneficial but not as beneficial as a history of no drinking or moderate drinking. Hence this attenuation factor, which can be used to conduct sensitivity analyses for further evaluation of this issue. In all subsequent analyses we used a downward adjustment of 20% to be on the conservative side.

Combining costs and effects

Once the treatment costs and the reductions in DALY disease burden have been computed for each scenario, it is a small step to also compute the difference between the costs of both health care systems as $\Delta(C) = C_1 - C_0$ and the difference between the effects as $\Delta(E) = E_1 - E_0$. The ratio, $\Delta(C)/\Delta(E)$, is the incremental cost-effectiveness ratio (ICER), which tells us whether the alternative health care system (current health care with additional online interventions) offers better value for money than the current health care system.

Handling uncertainty

As indicated, in probabilistic mode, ALCMOD takes parameter uncertainty into account. The uncertainty is captured by drawing values from the cost and effect distributions of all interventions at random and basing the calculations on these randomly drawn values. This can be repeated n times (in practice 500 times appears to be sufficient) and the outcomes of each of the iterations is stored in vectors of size n of the costs and effects of each of the scenarios, their differences, and the ICER. Following standard health economic modelling routines (Briggs et al 2006), the vectors are then used to produce ALCMOD's output, such as the mean and the median of the outcomes and several ICER plots and graphs.

Model assumptions

The assumptions of ALCMOD are shown in **textbox 3.2**.

Textbox 3.2. ALCMOD's assumptions and justifications

General assumptions:
- ALCMOD disregards the longer-term downstream costs, cost offsets, and health effects due to less drinking because the empirical literature rarely reports treatment effects beyond 12 months.
- Per-participant costs are assumed to follow a gamma distribution (Chisholm et al 2004b).
- Treatment effects, expressed in standardized mean difference scores, d, are assumed to follow the standard normal distribution, because d is almost equivalent to a z-score.
- The YLD (quality of life) differential is based on Sanderson et al.'s conversion factor (2004), which translates a change in disorder severity of size d induced by an intervention into a corresponding shift in the disability weight (DW) used in the YLD calculations.
- The YLL (mortality) differential is based on Gmel et al.'s (2003) relative risk of all-cause mortality stratified for level of pure ethanol intake (g/day).
- Costs and DALY outcomes have not been discounted because the focus is on short-term (<12 months) postimplementation (steady state) health and budget impacts.

Additional assumptions for the current simulations:
- The AUDIT distribution obtained from DrinkingLess is representative for the target population because this is a population of (former) problem drinkers still at risk of an alcohol use disorder and willing to seek treatment.
- Adherence rate is 50% for all interventions because a constant figure would help to obtain a clear view on cost-effectiveness due to fundamental changes in health care technologies.
- Alcohol intake is reduced by 20% after all interventions because the short-term contribution of YLL to the DALY disease burden is virtually negligible.
- All treatment effects on YLD and YLL are attenuated by 20% because the detrimental health effects of problem drinking are likely to linger on—even after return to moderate drinking or abstinence.

Results

Pre-intervention target group

In November 2009 we obtained data from the DrinkingLess monitoring system on 1,083 women and 2,538 men who participated in DrinkingLess. Their mean age was 44.7 years (SD 10.7). **Table 3.4** presents the observed AUDIT distribution for this population and the relative risk (RR) of premature death due to alcohol, as computed by ALCMOD.

Table 3.4: Pre-intervention characteristics of the target population

AUDIT score	Tentative label	Men % (N=987,000)	Women % (N=267,000)	Relative Risk (death) Men	Women
0-1	Abstinent	0.1	0.1	1.00	1.00
2-7	Moderate	1.6	3.6	0.86	0.96
8-15	Heavy	18.4	23.5	0.95	0.99
16-19	Hazardous	22.2	23.5	0.99	1.05
20-29	Harmful	50.1	43.4	1.10	1.12
30-40	Dependence	7.6	5.9	1.36	1.28

Comparing current care with new eHealth interventions added

We begin by comparing the current health care system (base case scenario) with an alternative scenario where eHealth interventions are added to conventional care. In this comparison, it was assumed that the new eHealth interventions would attract a different segment from the target population—a segment that would otherwise not have been the recipient of conventional care. Making this (unrealistic) assumption is a conscious choice, and we will return to it in the "Discussion" section. The results are as follows. The total health care costs in the base case scenario are €233 million. Adding new eHealth interventions would raise the health care expenditure to €319 million, an increase of €86 million. Under the base case scenario, 5,022 DALYs are averted; under the new scenario, this is doubled to 10,319 averted DALYs, an additional 5,296 averted DALYs (including 32 alcohol-related deaths that are avoided under the new scenario). Thus, the alternative health care system delivers more population health albeit at higher costs. **Figure 3.1** provides a corresponding visualization: the scatter of simulated costs and effects (due to uncertainty in the input parameters) corresponding to the alternative scenario is placed more to the north (more costs) and more to the east (more health) than the scatter belonging to the current health care system. Now, investing €86 million for averting 5,296 extra DALYs (i.e. €16,053/DALY) raises the question whether that would be money spent wisely. In the Netherlands, the willingness-to-pay for one averted DALY is about

€80,000 with a lower bound of €50,000. An even more conservative willingness-to-pay ceiling is customarily set at €20,000/DALY for non-fatal and mild disorders. It follows that the estimated €16,053/DALY falls well below any of the usual willingness-to-pay ceilings.

Figure 3.1: Total costs and effects in millions of euro (base case scenario versus the base case with eHealth interventions added)
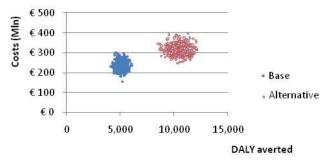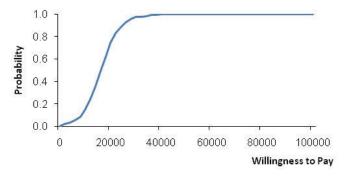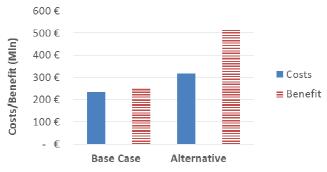

Figure 3.2, the ICER acceptability curve, represents a slightly different approach to the same issue. It depicts the probability that we must conclude that the new health care system is more cost-effective than the current system (vertical axis) for a range of willingness-to-pay ceilings (horizontal axis). For the simulated data, **figure 3.2** shows that the likelihood that the new health care system must be regarded as cost-effective increases sharply with increasing willingness-to-pay ceilings: the probability equals 0% when the willingness-to-pay for an additional health gain of one DALY averted is €0, increases to 50% at €16,000 and to 75% at €20,000. Beyond €30,000 the probability approaches certainty, and the conclusion that we must regard the new system as more cost-effective is no longer affected by higher willingness–to-pay levels. Again, accepting the threshold of €20,000/DALY implies that the new health system compares favourably with the current system in terms of cost-effectiveness.

ICER acceptability curve (base case scenario versus base case with eHealth interventions added)

Assuming for a moment that the willingness-to-pay for averting one DALY is €50,000, then we could directly compare the costs of the health care system (in euro) with health gains (also expressed in euro) by multiplying the averted DALYs by €50,000. **Figure 3.3**, a cost/benefit chart, shows ALCMOD's simulation results. The chart shows that costs and benefits are just balanced under the current health care system, while the benefits clearly outweigh the costs under the new system. To be more precise, every euro spent under the current system returns a value of about the same size (€1.08, i.e., a "surplus" of 8 euro cents), while the new health care system offers much better returns on investment: every euro spent, generates €1.62 in health-related value.

Figure 3.3: Cost-benefit chart in millions of euro (base case scenario versus base case with eHealth interventions added)
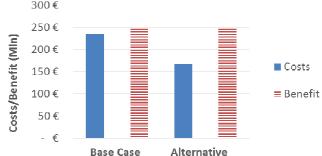

To summarize, the new health care system, with eHealth interventions added, is associated with higher health care delivery costs overall, but it would be a health care system which is more efficient than the current one, offering better value for money.

Comparing current care with new eHealth interventions with partial substitution

We also simulated another comparison, this one between a base case scenario that represents the current system and an alternative scenario with eHealth interventions added. In this case, however, the conventional face-to-face interventions are partly substituted by the new eHealth interventions. In this scenario, the coverage rate remains the same before and after the introduction of the eHealth interventions. Such a situation would arise if the eHealth interventions were to tap into the same target population seeking professional help, whether face-to-face, eHealth, or otherwise. In this scenario, the number of people who receive health care remains the same before and after the introduction of the new health technologies, and interventions are competing for the same target population and therefore partially substitute each other. In this scenario, ALCMOD computes that the number of DALYs averted under both systems is virtually the same: 4,984 DALYs under the current system and (exactly) 5,000 DALYs under the new system. In other words, partial substitution of conventional face-to-face interventions by eHealth interventions does not have any appreciable impact on population health. However, the overall cost of the new system is much lower at €166 million than the cost of the current system of €234 million, resulting in a cost saving of €68 million. **Figure 3.4** relays the same information. Again assuming a willingness-to-pay of €50,000/DALY, the cost-benefit ratio indicates that for every euro invested the generated health revenues are worth €1.06 (i.e., 6 euro cents surplus for every euro invested) under the current health care system. This improves to 52 euro cents surplus for every euro invested under the new scenario where the face-to-face interventions have been partly substituted by new eHealth interventions.

Figure 3.4. Cost-benefit chart in millions of euro (base case scenario versus partial substitution scenario)

Discussion

Main findings

The main rationale for introducing eHealth technologies is to increase timely access to health services, to reduce the costs of delivering health care, and to make more efficient use of the health care workforce. Indeed, ALCMOD's simulation results suggest that widespread implementation of eHealth interventions for alcohol use disorders would help to substantially increase population health in the Netherlands, albeit at higher system costs, when eHealth interventions are added to the existing health care system and more people become the recipients of the expanded system. The cost-effectiveness of the Dutch health system would also substantially improve if the new eHealth interventions were partially replacing some of the current face-to-face interventions. Then, adding eHealth interventions becomes a cost-effective option, because it will produce the same level of population health for a significantly smaller health care budget. The "truth" might be found somewhere between both extremes, because it is unlikely that the new eHealth interventions will exclusively recruit people that would otherwise not have been the recipients of conventional health care (as assumed in the first comparison), while it is also unlikely that the new eHealth interventions will tap into exactly the same pool of health care users (as assumed in second comparison). At any rate, both extreme scenarios carry the message that widespread introduction of eHealth technologies would help to substantially increase the efficiency of the Dutch health care system overall, with a more favourable cost-benefit ratio either way.

Strengths and limitations of ALCMOD

One of the benefits of a simulation model is that it helps to organize vast fields of knowledge across several disciplines. In the case of ALCMOD, these disciplines encompass addiction epidemiology and health economics, while the evidence that supports effect parameters is drawn from randomized clinical trials, meta-analyses, and evidence-based clinical guidelines. In addition, a model makes all the necessary information available in a dynamic form, permitting 'what-if' analyses. This could be of assistance to policy formulation. ALCMOD is therefore best seen as a decision-making support tool, capable of giving almost instant feedback on policy-makers' attempts to find an optimal solution in the context of constrained decision-making in a complex environment. ALCMOD can also be employed for setting research agendas. After all, it helps to identify those parameters (within the set of parameters included in the model) that have an impact on health gains and costs.

When some of these parameters are surrounded by a non-negligible degree of uncertainty (thereby causing unwanted uncertainty in the decision-making process), then empirical research is recommended, with the aim of reducing uncertainty in those parameters. Furthermore, ALCMOD can assist in identifying opportunities for system innovation by simulating hypothetical interventions, for example, an adjunctive intervention that helps to enhance treatment adherence. Among other strengths of ALCMOD are its adaptability to other countries, settings, and target groups and its capability to explicitly model treatment coverage and adherence rates. Finally, ALCMOD conducts automated multivariate uncertainty analyses to quantify uncertainty in costs, effects, and related parameters.

ALCMOD is subject to several limitations that need to be taken into account when interpreting ALCMOD's outcomes. First, ALCMOD's outcomes are modelled as *steady-state* population averages, and it is not clear when a health care system finds equilibrium after the introduction of new health technologies. This is unlikely to occur instantaneously and might take as long as several years. Second, it should be borne in mind that the introduction of new health technologies entails costs of its own, but the costs of introducing new technologies are not incorporated in ALCMOD's output. In fact, ALCMOD's output captures only the costs of offering a package of interventions once the interventions have been fully implemented. However, it will always take effort, time, and expenditure before the results of an improved health care system become visible in real life. Third, introduction of eHealth technologies may have unforeseen consequences that may increase longer-term health care costs, for example, by supply-induced demand for health care, thus attracting people to the health care system who otherwise would not have become dependent on (expensive, face-to-face delivered) health care. Fourth, it should be understood that ALCMOD focuses on short-term health impacts. Thus, ALCMOD ignores the longer-term impacts on quality of life, mortality, and health care utilization and it should be understood that longer-term impacts depend, in part, on a wide range of alcohol-related disorders that usually occur later in life. Since these longer-term effects are mainly related to the more severe alcohol use disorders, ALCMOD is unlikely to capture the full benefits of interventions for the severe disorders and may thus give undue weight to the less severe disorders. Fifth, ALCMOD is limited in that it only models clinical interventions while disregarding other alcohol-control options, such as banning alcohol advertising, taxing, restricting access to alcoholic beverages, and improving road safety, although these nonclinical interventions are likely to be (very) cost-effective (Chisholm et al 2004b). In the same vein, ALCMOD regards only the cost impacts incurred by the health care system, while disregarding costs and cost-

offsets outside the health care system, such as patients' out-of-pocket payments to access services, changes in labour productivity, and costs incurred by the criminal justice system. To summarize ALCMOD's basic assumptions: ALCMOD only models incremental health gains and health care delivery costs over the shorter time horizon, assuming a steady state in the modelled health care systems. See **textbox 3.2** for a summary of ALCMOD's assumptions and their justifications.

Conclusion

It is not immediately clear if our findings are valid for countries other than the Netherlands. After all, in low-income countries, labour might be less costly than the capital inputs required for the new eHealth technologies. Also the population's access to the Internet could be an issue. Moreover, one could encounter cultural obstacles to using the Internet for alcohol use disorders. Such factors might impinge on coverage and adherence rates and mitigate impacts on population health, ultimately diminishing the cost-effectiveness of new health technologies.

To illustrate, in the Netherlands, close to 90% of the population has access to the Internet, and Internet usage is distributed fairly evenly across demographic groups, but in other countries, Internet usage might be concentrated in only some population segments. In addition, it is worth noting that the emergence of mobile technologies may offer an opportunity to offer eHealth interventions for population segments that otherwise might be hard to reach. Therefore, the question as to whether eHealth will deliver the same benefits to other countries is best addressed per country, per setting, and per target group. Ante hoc assessment of the cost-effectiveness of innovations in health care systems may help to inform policy decisions. ALCMOD was created with exactly these aims in mind.

We recommend that ALCMOD be used in an iterative consensus building process that encompasses all pertinent stakeholders (e.g., health care users, health care providers, health care financiers, and health policy planners) who can review and make amendments to modelled scenarios. Recently, we had an encouraging experience with such an approach while using a similar model for the treatment of depressive disorder. In any case, we would advise against using ALCMOD as an autopilot for policymaking. After all, setting priorities for health care delivery is about acceptability and equity, as well as about cost-effectiveness considerations. As always, we need to base decisions on the best judgments and evidence available, but the evidence that informed ALCMOD points toward the conclusion that eHealth

interventions can help to bridge the mental health gap by bringing scalable and cost-effective health services within reach of all who have access to the Internet—literally at their fingertips.

Improving the cost-effectiveness of a health care system for depressive disorders by implementing telemedicine: a health economic modelling study

Abstract

Objectives: Depressive disorders are significant causes of disease burden and are associated with substantial economic costs. It is therefore important to design a health care system that can effectively manage depression at sustainable costs. This article computes the benefit-to-cost ratio of the current Dutch health care system for depression, and investigates whether offering more online preventive interventions improves the cost-effectiveness overall.

Methods: A health economic (Markov) model was used to synthesize clinical and economic evidence and to compute population-level costs and effects of interventions. The model compared a base case scenario without preventive telemedicine and alternative scenarios with preventive telemedicine. The central outcome was the benefit-to-cost ratio, also known as return-on-investment (ROI).

Results: In terms of ROI, a health care system with preventive telemedicine for depressive disorders offers better value for money than a health care system without Internet-based prevention. Overall, the ROI increases from €1.30 ($1.55) in the 'no prevention' base case scenario to €1.61 ($1.92) in the 'prevention-only' alternative scenario in which preventive telemedicine is offered. In the scenario where prevention is added to the current intervention mix, the ROI increases to €1.32 ($1.58). In the scenario in which the costs of offering preventive telemedicine are balanced by reducing the expenditure for curative interventions, ROI is also €1.32 ($1.58), but while keeping the health care budget constant. Lowering the coverage of the relatively cost-inefficient curative interventions does not further increase the ROI as this is only a minor change in terms of reduced coverage.

Conclusions: For a health care system for depressive disorders to remain economically sustainable, its cost-benefit ratio needs to be improved. Offering preventive telemedicine at a large scale is likely to introduce such an improvement.

Introduction

Depression is a leading cause of non-fatal disease burden (Ustun et al 2004; Mathers and Loncar 2006; Saarni et al 2007; Vos et al 2012) and has substantial economic consequences (Berto et al 2000; Greenberg et al 2005; Smit et al 2006c; Vasiliadis et al 2013). Reducing the disease burden due to depressive disorders at affordable costs is therefore of great significance to public health.

Cushioning the adverse effects of depression requires a health care system well equipped to manage the disorder. To that end, the interventions for depression that are offered need to be acceptable to both health care users and health care providers. In addition, the interventions must be effective in generating the required health gains and be economically sustainable over time. It is difficult to identify which particular combination of interventions will meet all these criteria within the extensive range of available options that are offered in multiple formats to different target groups.

The task of identifying an "optimal" health care system becomes even more daunting when the acceptability and cost-effectiveness of a newly designed health care system have to be compared with the cost-effectiveness of the current health care system. In particular, we need to know how a (hypothetical) health care system based on widespread implementation of preventive telemedicine would compare with the current health care regimen without preventive telemedicine. Would such a health care system produce larger health gains? In addition, how would the new system compare with the current health care regimen in terms of its benefit-to-cost ratio?

To facilitate decision making, we developed a health economic simulation model for depression called DEPMOD. This model assesses the population-level cost-benefit ratio of an alternative health care system relative to the current one. Although availability of data prompted us to apply DEPMOD to the population aged 18-65 years, we expect that DEPMOD is also relevant to older populations. This is especially true because the older population has an elevated risk for depression (Licht-Strunk et al 2007) and the evidence suggests an increased risk of additional adverse outcomes for older people with depression (Byers et al 2012). The older population might be under pressure to be economically productive, even beyond the current age of retirement, due to the present-day economic downturn in "greying" societies.

At the same time, increased life expectancy, common in high-income countries, is associated with an increase in the number of depressed older people.

In sum, greying societies, increased demand for mental health care, rising health care expenditure, and dwindling labour forces for mental health underscore the importance of the health care system being reassessed and geared toward offering more cost-effective interventions. Implementing interventions that can be offered over the Internet seems to be a promising approach because these interventions are likely to be scalable, effective, and cost-effective. DEPMOD simulates the possible consequences of offering Internet interventions for major depression.

Experience with the Australian Assessing Cost-Effectiveness models for heart disease, mental disorders, and prevention (Vos et al 2005; Mihalopoulos et al 2011b) and the WHO-CHOICE models (CHOosing Interventions That Are Cost-Effective) (Hutubessy et al 2003; Chisholm et al 2004a) indicates that health economic models may have value for policymaking. DEPMOD was specifically designed for the Dutch health care system, using Dutch population-based cohort data on depressive disorder (Bijl et al 2002) and standard cost prices pertinent to the Dutch health care system (Hakkaart-van Roijen et al 2010). It also models the impact of several preventive e-health interventions that were recently developed, evaluated, and disseminated in the Netherlands. However, DEPMOD can be used for other countries and populations, provided that data requirements are met.

The aim of the current article was to briefly describe DEPMOD and then apply DEPMOD by modelling the current package of health care interventions and an extended package in which preventive telemedicine is added. The goal was to address the question of whether preventive telemedicine offers good value for money. To this end, we report the 'return-on-investment' (ROI), which expresses the euro (dollar) value of (health-related) benefits associated with every euro (dollar) invested in preventive telemedicine. This may guide decision-making in health care policies.

We define telemedicine (e-health) as psychological self-help interventions that are delivered over the Internet, either with or without minimal therapist support. Meta-analyses of randomized trials have demonstrated the effectiveness of both prevention of depressive disorder (Cuijpers et al 2008a; Muñoz et al 2010; van Zoonen et al 2014) and (preventive) e-health interventions (Spek et al 2007a; Andersson et al 2009). In addition, telemedicine is very scalable because of the widespread usage of the Internet. It should be noted that older people are the fastest

growing group of new Internet users, and one of the main reasons older individuals use the Internet is because they are seeking answers to health questions. By implication, there is a good match between older people's Internet usage and e-mental health. Although not explicitly modelled here, evidence suggests that depression prevention is also effective in the older population (van 't Veer-Tazelaar and van Marwijk 2011). The goal of the current article was to synthesize the relevant clinical and economic evidence in a health economic modelling study.

METHODS

Comparing scenarios: usual care versus more preventive telemedicine
DEPMOD is used to compute the cost-benefit ratio by comparing "usual care" with an alternative scenario in which usual care is augmented with preventive telemedicine (Scenario A). Scenario B compares usual care with a 'prevention-only' scenario. In addition, Scenario C is analysed in which the costs of offering additional preventive telemedicine are compensated for by reducing the health care budget for curative interventions, thereby keeping the overall costs of the new scenario under the current budgetary ceiling.

The usual care scenario, which forms the basis for the comparisons, is an evidence-based health care system that is fully in agreement with the Dutch clinical guidelines for the treatment of depression. Because it is likely to be better than the current Dutch health care system, we refer to it as "enhanced usual care" (**table 4.1**). This long list of evidence-based interventions was then used to select only those interventions that were acceptable from a patient's point of view and were appropriate from a health care professional's point of view. To that end, focus groups were used; a panel of 17 health care users judged to what extent they would be willing to accept and actively engage in each of the interventions, whereas a panel of 10 health care professionals judged to what extent the interventions were appropriate to offer for the various manifestations of depressive disorder. Both panels showed a relatively high degree of consensus with regard to their preferences (Cronbach alpha = 0.79 for care users and 0.70 for care providers). Taking these preferences into account, the extensive evidence-based interventions was reduced to a shorter list of interventions that are not only evidence-based but also preference-based (**table 4.1**).

The list of evidence-based and preference-based interventions forms the basis for performing scenario analysis and is likely to be more cost-effective than usual

care. In usual care, not every intervention is evidence-based or meets with approval by both care users and health care providers.

Table 4.1: Selected evidence-based and preference-based interventions by depression severity level: Costs (in 2009 euro and dollar), Compliance with therapy (%) and Effect, as risk difference (RD) or as standardized effect size, d, all representing average values

Intervention by depression severity level	Costs	Compliance rate	Effect
Subclinical depression	*€ ($)*	*%*	*RD*
Self-help book[1]	*348(413)*	*52*	*0.062*
Group course: 8-10 sessions[2]	*506(601)*	*64*	*0.055*
E-health intervention (unsupported) [3]	*178(211)*	*56*	*0.077*
Mild depression	*€ ($)*	*%*	*d*
E-health intervention (supported) [4]	*313(372)*	*43*	*0.32*
"Interapy": online psychotherapy, 10 sessions CBT[5]	*2,154(2,558)*	*44*	*0.70*
Individual psychotherapy, primary care, 8 sessions[6]	*1,296(1,539)*	*56*	*0.69*
Moderate depression	*€ ($)*	*%*	*d*
E-health intervention (supported) [7]	*313(372)*	*43*	*0.32*
"Interapy": online psychotherapy, 14 sessions CBT[8]	*2,154(2,558)*	*44*	*0.70*
Individual psychotherapy, primary care, 8 sessions[9]	*1,296(1,539)*	*56*	*0.69*
Severe depression	*€ ($)*	*%*	*d*
Individual psychotherapy, outpatient care, 8-24 sessions[10]	*1,447(1,719)*	*68*	*0.70*
Anti-depressants, 3-6 months via GP[11]	*235(279)*	*44*	*0.72*
Anti-depressants, 3-6 months, with additional psychological support[12]	*289(343)*	*56*	*0.72*
Combination therapy (medication and psychotherapy) [13]	*1,215(1,443)*	*65*	*1.05*
Recurrent depression	*€ ($)*	*%*	*RR*
Clinical management with maintenance medication, 12 months[14]	*537(638)*	*42*	*0.75*
Preventive Cognitive Therapy: 8 group sessions[15]	*406(482)*	*63*	*0.73*
Supported self-help PCT: via the internet[16]	*403(479)*	*46*	*0.73*

1. *Taken from Willemse et al 2004*
2. *Taken from Cuijpers 1998, Cuijpers et al 2008a*
3. *Taken from van Zoonen et al 2014*
4. *Taken from Spek et al 2007a, Cuijpers et al 2008b*
5. *Taken from Ruwaard et al 2009*
6. *Taken from Cuijpers et al 2007, 2008c*
7. *Taken from Spek et al 2007a; Cuijpers et al 2008b*
8. *Taken from Ruwaard et al 2009*
9. *Taken from Cuijpers et al 2007; 2008c*
10. *Taken from Ekers et al 2008*
11. *Taken from Aroll et al 2005, Kirsch et al 2008, Fournier et al 2010*
12. *Taken from Kirsch et al 2008; Fournier et al 2010*
13. *Taken from Cuijpers et al 2009*
14. *Our reanalysis of the meta-analysis by Vittengl et al. (2007)*
15. *Our reanalysis of the meta-analysis by Vittengl et al. (2007)*
16. *See 15. Hypothetical effect size on the assumption that supported e-health is as effective as face-to-face delivered prevention of recurrence, albeit associated with a lower adherence rate*

Table 4.2 describes the scenarios that were analysed by using DEPMOD. First, the base case scenario of evidence-based and preference-based care without prevention was assessed, in which coverage rates and adherence rates were elicited from the focus groups. The alternative scenario (Scenario A) is essentially the same as the base case scenario, except prevention and (preventive) telemedicine is now offered. To be more specific, prevention consists of face-to-face interventions with an arbitrarily low coverage rate set at 2%. Preventive e-health interventions are offered at a coverage rate of 15%, which is likely to be attainable in practice (Riper et al 2007). E-health interventions for prevention of relapse and recurrence are assumed to be somewhat lower, with coverage set at 10%. In Scenario B, a prevention-only scenario was analysed in order to determine the unique contribution of prevention on cost-effectiveness. Finally, Scenario C is like Scenario A, but with treatment interventions scaled down, thus keeping the overall costs balanced (equal to the costs in the base case scenario). Coverage rates in the alternative scenarios are hypothetical and can be used to conduct 'what-if' analyses around potentially interesting health care systems. The remainder of the methods section describes DEPMOD, which is based on methods as described by Briggs et al. (2006) and Drummond et al. (2005).

Table 4.2: Modelled scenarios: coverage rates (%) for each of the interventions by depression severity level

Intervention by depression severity level	Base case	Alt. A	Alt. B	Alt. C
Subclinical depression	**%**	**%**	**%**	**%**
Self-help book	0	2	2	2
Group course: 8-10 sessions	0	2	2	2
E-health intervention (unsupported)	0	15	15	15
Mild depression	**%**	**%**	**%**	**%**
E-health intervention (supported)	2	2	0	1.9
"Interapy": online psychotherapy, 10 sessions CBT	2	2	0	1.9
Individual psychotherapy, primary care, 8 sessions	17	17	0	15.8
Moderate depression	**%**	**%**	**%**	**%**
E-health intervention (supported)	2	2	0	1.9
"Interapy": online psychotherapy, 14 sessions CBT	2	2	0	1.9
Individual psychotherapy, primary care, 8 sessions	16	16	0	14.9
Severe depression	**%**	**%**	**%**	**%**
Individual psychotherapy, outpatient care, 8-24 sessions	18	18	0	16.7
Anti-depressants, 3-6 months via GP	20	20	0	18.6
Anti-depressants, 3-6 months, with additional psychological support	20	20	0	18.6
Combination therapy (medication and psychotherapy)	16	16	0	14.9
Recurrent depression	**%**	**%**	**%**	**%**
Clinical management with maintenance medication, 12 months	0	2	2	2
Mindfulness-based PCT: 8 group sessions	0	2	2	2
Supported self-help PCT: via the internet	0	10	10	10

DEPMOD

Conceptually, DEPMOD models the epidemiology of major depression and simulates how a health care system affects the incidence (via prevention), prevalence (via treatment), and recurrence (via relapse prevention) of the disorder. Generating health impacts by offering interventions entails costs. Both the costs and the health gains are evaluated by using DEPMOD. The epidemiology of depression is modelled as a series of transitions between different health states (healthy, depressed, and death), taking into account both severity of depression (subclinical, mild, moderate, and severe depression) and the number of depressive episodes (recurrences). The simulated health care system consists of a mix of preventive interventions, curative interventions (for mild, moderate, and severe depression), and interventions to prevent recurrences, as outlined in **tables 4.1** and **4.2**.

The purpose of DEPMOD is to calculate the total health care expenditure and health gains under the current health care system, and to compare the current scenario with the alternative scenarios. The following sections describe the model, the data, and the underlying assumptions in more detail.

Epidemiology

DEPMOD is restricted to depressive disorder, as defined according to the Diagnostic and Statistical Manual of Mental Disorders, Third Edition (American Psychiatric Association 1987). DEPMOD assumes a population of 10 million people, aged 18-65 years. Estimates of incidence (158,000 new cases per year), episode duration (6 months on average), prevalence (588,600 acute cases annually), and recurrence rates of depressive disorder (45% of the currently depressed people have a history of previous episodes) were obtained from The Netherlands Mental Health Survey and Incidence Study, a population-based psychiatric epidemiologic cohort study (Bijl et al 2002; de Graaf et al 2012b). Depression-specific mortality rates were assessed by using a meta-analytic approach (Cuijpers and Smit 2002). DEPMOD takes into account that the risk of yet another depressive episode increases with the number of previous episodes (after Le Lay et al 2006).

Health care system

A health care system consists of preventive interventions to reduce incidence; treatment of mild, moderate, and severe depression to reduce disease burden; and relapse prevention in recovered patients to reduce risk of relapse and recurrence. These factors of primary prevention, cure, and relapse prevention can be considered

a system of health care "echelons" along the disease continuum. Each echelon consists of a mix of interventions.

Each intervention is described by its impact on health (Cohen's d, RD or RR), coverage rate (percentage of population receiving the intervention), adherence rate (extent to which patients comply with the intervention), and cost (per intervention per patient). Effects were based on meta-analyses where possible, and randomized controlled trials or estimates otherwise (**table 4.1**). Costs were estimated by mapping the amount of time of health care professionals per intervention multiplied by hourly rates.

The sum of all cost and total health gains were calculated at the level of the population. Costs were restricted to direct medical cost (in euro for the reference year 2009, converted to USD by using purchasing power parities) (OECD stats accessed June 1, 2014). Unit cost prices were obtained from the Dutch Guideline for Health Economic Evaluations (Hakkaart-van Roijen et al 2010). Health gains are expressed as a reduction in the disease burden due to depression (i.e., fewer disability-adjusted life-years [DALYs]).

Assessing health gains

Health care interventions aim to reduce the number of DALYs in the population. DALY is a measure of disease burden in a population, taking into account two components of disease burden: morbidity and mortality. Morbidity is related to time spent in a health state characterized by a lowered quality of life due to disability. Mortality comes into the equation when illness is associated with premature death. Murray et al. (1997) present a description of DALYs.

In DEPMOD, DALY reductions are achieved in two ways: by preventing people from becoming depressed through primary prevention and by treating people who have depression, thereby lowering their disease burden. DALY reductions due to treatment are based on the effect size d associated with treatment, in addition to using Sanderson's conversion method (Sanderson et al 2004). DALY reductions due to prevention arise as people are prevented from entering a depressed stated, which is associated with morbidity and excess mortality.

Cost-effectiveness analysis

To allow for parameter uncertainty in costs and effects, the model randomly draws a value from the distributions assigned to the parameters and computes the outcome for that configuration of parameter values. This procedure is repeated 1,000 times over all parameters simultaneously. In each run, the outcomes (costs and health gains

for each scenarios) are computed and stored in DEPMOD's memory. Then, following the methods of Briggs et al. (2006), all 1,000 simulated outcomes are evaluated simultaneously, thus explicitly accounting for uncertainty in the input parameters.

After generating 1,000 values of costs and DALYs for the current and alternative health care systems, costs and effects are discounted when the time horizon exceeds 1 year (the analysis presented is based on a time horizon of five years). Discounting rates (1.5% for the effects and 4.0% for the costs, per the pertinent economic guideline (Hakkaart-van Roijen et al 2010)) are automatically subjected to further sensitivity analyses. In the next step, differences in the costs (incremental costs) and differences in DALYs (incremental effects) across both scenarios are obtained, and an estimate of the incremental cost-effectiveness ratio (ICER) is computed: ICER = $(C_1 - C_0)/(E_1 - E_0)$, where C are costs, E are effects, and subscripts 1 and 0 refer to the alternative and base case scenarios, respectively. The ICER is one of the key outcomes of an economic evaluation (see Briggs et al 2006).

Our time horizon was five years, but this could be changed to a minimum of one year. Finally, the return on investment (ROI) of each scenario is calculated by dividing DALY health gains, conservatively valued at €20,000 ($23,860) per DALY, by total cost.

In health economic modelling, making assumptions is inevitable. Whenever assumptions were made, we used conservative estimates to decrease the risk of outcomes being overly optimistic. It is important to understand how the assumptions affect the outcome of the model. **Textbox 4.1** presents DEPMOD's main assumptions, their justifications, and their possible impact on the findings.

Textbox 4.1. DEPMOD's assumptions, justifications and implications

Assumption	Justification	Implication
The 1-year incidence is constant at 158,000 cases per year. Prevalence is 588,600 annually in the adult Dutch population of 10 million people.	Data obtained from The Netherlands Mental Health Survey and Incidence Study, a population-based, psychiatric epidemiologic cohort study (Bijl et al 2002; de Graaf et al 2012b).	Prevalence determines the cost and effects. The ratio incidence/ prevalence determines the relative importance of prevention.
Episode duration is 6 months on average.	After Spijker et al. (2002).	Taking episode duration into account affects health benefits. A shorter duration means less potential to generate health benefits.
It is possible to have up to five recurring episodes of depression. After the fifth recurrence, a patient is assumed to be chronically depressed. Recurrence rates of depressive disorder are 50%, 70%, 80%, 85%, and 90% for the first to the fifth episode.	Relapse rates are higher after a previous depressive episode.	Increasing risk of recurrence results in patients making heavier demands on the health care system, which emphasizes the importance of preventing recurrence from a cost-effectiveness point of view.
Effects are normally distributed.	After Briggs et al. (2006).	Uncertainty around the effect parameters is symmetrical.
Costs are gamma distributed.	After Briggs et al. (2006).	Uncertainty around the cost parameters is skewed to the right.
WTP for averting one DALY is €20,000 ($23,755).	WTP values for averting one DALY often range between €20,000 ($23,755)and €80,000 ($95,020) (Council for Public Health and Health Care 2006). A relatively low number of €20,000 ($23,755) was chosen to be conservative.	A health care system is deemed cost-effective when the price per one DALY averted is less than the WTP ceiling of €20,000 ($23,755).

Effects of CBT are maintained over at least 1 year after treatment, but effects of pharmaceutical interventions decline almost instantly after discontinuation.	Based on analysis after Willemse et al. (2004).	Longer lasting prophylactic effects for CBT than for pharmaceutical interventions amounts to an increased cost-effectiveness of CBT relative to antidepressant medication.
CBT offered during the acute stage of depression introduces a prophylactic effect (which, in our model, prevents relapse in ±1.5-5% of patients with mild, moderate or severe depression).	After Willemse et al. (2004).	The presence of a prophylactic effect makes it more desirable to treat acute cases of depression with CBT because it may help to avoid new onsets of the disorder in the future.
CBT: cognitive behavioural therapy; DALY: disability-adjusted life-year; WTP: willingness-to-pay		

RESULTS
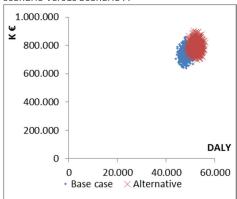
Alternative A
The first comparison (the base case scenario versus the alternative Scenario A) evaluates the added value of offering preventive interventions in terms of improvement in the cost-benefit ratio of the health care system. Cost and effects were modelled out over a period of five years. We present the key findings.

First, a health care system with indicated prevention and relapse prevention costs 7% more than a system without preventive telemedicine. Second, health gains are 9% higher in the scenario with preventive e-health. Third, in an evidence-based and preference-based system without preventive e-health, a mean (standard deviation [SD]) cost of €15,446 [€813] ($18,427 [$970]) is required for averting one DALY of disease burden. Costs per averted DALY drop to a slightly more favourable €15,163 [€751] ($18,090 [$896]) when e-health is offered. This means that the costs for averting one DALY decline slightly as a result of adding web-based prevention, illustrating that the health care system in its entirety becomes more cost-effective, even though costs increase due to additional investments in preventive telemedicine. Finally, when averting one DALY is economically valued at €20,000 ($23,860), the ROI in the base scenario (without prevention) amounts to €1.30 [€0.07] per euro invested

in health care ($1.55 [$0.08] per dollar invested). The ROI improves when prevention is added (alternative A) and becomes €1.32 [€0.07] ($1.58 [$0.08]).

In sum, the data suggest that offering preventive telemedicine makes the health care system more cost-effective, even though offering preventive telemedicine introduces costs of its own. However, the impact on cost-effectiveness when adding prevention is damped, as curative interventions still have a large share in the overall health care system in scenario A. **Figure 4.1** shows that adding prevention causes a modest shift in the overall costs and effects.

Figure 4.1. Simulation output costs and disability-adjusted life year (DALY) averted in the base case scenario versus Scenario A

Alternative B
In order to determine the impact of prevention, we compare the base case scenario with the alternative 'prevention-only' scenario B. Cost and effects were again modelled out over a period of five years.
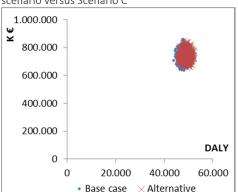
In the prevention-only scenario, costs per averted DALY equal €12,505 [€1,061] ($14,919 [$1,266]), which is lower than the costs per DALY of €15,446 [$18,427] in the scenario of enhanced usual care. When averting one DALY is economically valued at €20,000 ($23,755), the ROI in the prevention-only scenario equals €1.61 [€0.14] ($.1.92 [$0.16]), which is significantly higher than the ROI associated with enhanced usual care.

The prevention-only scenario suggests that offering preventive telemedicine makes the health care system more cost-effective, even though offering preventive

telemedicine introduces costs of its own. This impact becomes smaller as the relative share of preventive interventions (as compared to curative interventions) gets smaller.

Alternative C

The next scenario introduces the same increase in preventive telemedicine but a decrease in the coverage of (curative) interventions offered in the base scenario by 7% to keep the total cost of the health care system balanced. Again, the alternative scenario is compared with the base case scenario and is modelled out over a period of five years. Findings are as before, yet slightly more favourable. First, because of the decreased treatment costs in the alternative scenario, total costs do not change. Second, due to the relative cost-effectiveness of preventive e-health, health gains increase by 2%. Third, as before, it costs (mean [SD]) €15,446 [€813] ($18,427 [$970]) to reduce the disease burden of depression by one DALY in an evidence- based and preference-based system without preventive e-health. Under the alternative Scenario C, this amount becomes €15,145 [€746] ($18,068 [$891]) per averted DALY. Finally, following the same line of reasoning, the ROI increases to €1.32 [€0.07] ($1.58 [$0.08]) in alternative C.

The corollary is that offering preventive e-health interventions makes the health care system more cost-effective because a larger health gain is achieved while keeping costs equal. **Figure 4.2** demonstrates that the DALY gains in Scenario C are slightly higher, while costs in both scenarios are comparable. However, as prevention constitutes only a relatively small part of the health care system, the overall impact is small (as noted by the overlapping uncertainty intervals in **figure 4.2**).

Figure 4.2. Simulation output costs and disability-adjusted life year (DALY) averted in the base case scenario versus Scenario C

CONCLUSIONS

Main findings

The main finding of the current study was that e-health interventions which seek to prevent onset of first and later episodes of depression can help to make the health care system for depressed patients more cost-effective overall. Thus, a health care system for depressive disorders that is both evidence-based and preference-based (i.e., evidence-based interventions that are met with approval of both health care users and health care providers) represents a good ROI. Modelled out over a period of five years, every euro (dollar) spent would generate health gains worth €1.30 ($1.55), assuming that averting one DALY is conservatively valued at €20,000 ($23,860). However, the same health care system with realistic levels of preventive telemedicine implemented, and fewer curative interventions, would produce a slightly better return of €1.32 ($1.58) of health-related value for every euro (dollar), as prevention-only is associated with a relatively high return of €1.61 ($1.92) for every euro (dollar).

Although the model is based on a population aged 18-65 years, we believe comparable results are likely to be obtained for older populations. Evidence suggests that offering telemedicine to older people is promising. In a review on telecare for elderly people with chronic diseases, patients were generally satisfied, accepted the technology, and enjoyed self-monitoring (Botsis and Hartvigsen 2008). In addition, evidence specifically on treating depression in older people with telemedicine is promising. E-health interventions proved to be effective in treating depressive

symptoms in older people (Spek et al 2007a; 2007b), and in a sample of mainly older people, telemedicine was successfully used to adapt a collaborative care model for depression (Fortney et al 2007). In addition, from a demographic perspective, the current generation represented by our data are the elderly of the future. We may have to substantially rely on health technologies in the future that are less labour intensive than our current health care models.

Strengths and limitations

One of the benefits of a simulation model is that it helps to organize vast fields of knowledge across several disciplines. In the case of DEPMOD, these disciplines encompass psychiatric epidemiology and health economics, while the evidence that supports effect parameters is drawn from randomized clinical trials, meta-analyses, and evidence-based clinical guidelines. It also proved possible to elicit patients' preferences for certain interventions and to incorporate these preferences into the model. The model makes all information available in a dynamic form, which makes it possible to conduct 'if-then' analyses. This could be of assistance when exploring options for health care policies.

Our study has a number of limitations that need to be acknowledged. In health economic modelling, much depends on the assumptions made in the model. Whenever we had to make an assumption, we tried to make a conservative one; that is, an assumption that is likely to portray a not overly optimistic outcome scenario. For example, we used the more conservative value of €20,000 ($23,860) for averting one DALY and not the more generous value of €50,000 ($59,650), which is frequently suggested in the literature. Although we accounted for parameter uncertainty to some extent by using extensive sensitivity analyses, we emphasize that the value of our model lies in the comparative analysis of different health care scenarios rather than the interpretation of absolute values.

Another limitation is that the model is based on a population aged 18-65 years. Data available on the population older than 65 years are relatively scant, although evidence seems to suggest that the older population is willing to use and is receptive to telemedicine interventions in general and depression-oriented telemedicine in particular (Fortney et al 2007; Spek et al 2007a; 2007b; Botsis and Hartvigsen 2008). Although we considered a population at working age, only health care costs were considered in the analysis. Future research should aim to map the cost impact of prevention of depression from the broader, societal perspective. Even though our model is based on Dutch data, DEPMOD can be used in other countries

as well. With the appropriate data on epidemiology, effectiveness of interventions, and costs, DEPMOD could be adapted to different contexts and population segments. Thus, diverse populations could be investigated by running DEPMOD separately for each population segment.

It should also be noted that implementing telemedicine on a large scale entails costs of its own. DEPMOD did not include the costs of making such a transition from one health care system to another. However, the model did compare the benefit-to-cost ratios of two health care systems after full implementation (i.e., when the systems were in a steady-state balance). It is worth noting that implementation, especially in the presence of a culturally diverse population, is challenging in its own right.

For these reasons, DEPMOD is best seen as an explorative decision support tool. It is able to give almost instant feedback on policymakers' attempts to select the economically more attractive scenario in the context of constrained decision making under uncertainty in a complex environment. We recommend that DEPMOD be used in an iterative consensus-building process that encompasses all pertinent stakeholders (e.g., health care users, health care providers, policymakers). In any case, we would advise against using DEPMOD as an autopilot for policymaking.

DEPMOD can also be used for setting research agendas because it helps to identify those parameters that have an impact on health gains and costs. If any of these parameters is surrounded by a nontrivial amount of uncertainty, it is recommended to conduct empirical research with the aim of reducing uncertainty in that parameter. Finally, we wish to emphasize that ante hoc modelling requires empirical validation later. It is thus recommended that studies be conducted to test the hypotheses suggested by the modelling study.

Implications

Our modelling work shows that preventive interventions, and especially preventive e-health interventions, have the potential to improve the cost-effectiveness of the health care system. This finding is consistent with other modelling studies on prevention (Mihalopoulos et al 2011a; van den Berg et al 2011) and e-health (Smit et al 2011). Given the rising demand for health care and the corresponding increase in health care expenditure, preventive telemedicine could play an important role, especially in greying societies in which access to the Internet is available to almost all citizens.

Part III

New approaches to improving the cost-effectiveness of health care

Chapter 5. Mental health care system optimization from a health-economics perspective: where to sow and where to reap?*

Chapter 6. Improving the cost-effectiveness of health care by simultaneously investing and disinvesting in interventions

* was granted the Excellence in mental health policy and economics research award 2015

Chapter 5

Mental health care system optimization from a health-economics perspective: where to sow and where to reap?

Abstract

Background: Health care expenditure (as % of GDP) has been rising in all OECD countries over the last decades. Now, in the context of the economic downturn, there is an even more pressing need to better guarantee the sustainability of health care systems. This requires that policymakers are informed about optimal allocation of budgets. We take the Dutch mental health system in the primary care setting as an example of new ways to approach optimal allocation.

Aims of the study: To demonstrate how health economic modelling can help in identifying opportunities to improve the Dutch mental health care system for patients presenting at their GP with symptoms of anxiety, stress, symptoms of depression, alcohol abuse/dependence, anxiety disorder or depressive disorder such that changes in the health care system have the biggest leverage in terms of improved cost-effectiveness. Investigating such scenarios may serve as a starting point for setting an agenda for innovative and sustainable health care policies.

Methods: A health economic simulation model was used to synthesize clinical and economic evidence. The model was populated with data from GPs' national register on the diagnosis, treatment, referral and prescription of their patients in the year 2009. A series of 'what-if' analyses was conducted to see what parameters (uptake, adherence, effectiveness and the costs of the interventions) are associated with the most substantial impact on the cost-effectiveness of the health care system overall.

Results: In terms of improving the overall cost-effectiveness of the primary mental health care system, substantial benefits could be derived from increasing uptake of psycho-education by GPs for patients presenting with stress and when low cost interventions are made available that help to increase the patients' compliance with pharmaceutical interventions, particularly in patients presenting with symptoms of anxiety. In terms of intervention costs, decreasing the costs of antidepressants is expected to yield the biggest impact on the cost-effectiveness of the primary mental health care system as a whole. These "target group – intervention" combinations are the most appealing candidates for system innovation from a cost-effectiveness point of view, but need to be carefully aligned with other considerations such as equity, acceptability, appropriateness, feasibility and strength of evidence.

Discussion and limitations: The study has some strengths and limitations. Cost-effectiveness analysis is performed using a health economic model that is based on registration data from a sample of GPs, but assumptions had to be made on how these data could be extrapolated to all GPs. Parameters on compliance rates were obtained from a focus group or were based on mere assumptions, while the clinical

effectiveness of interventions were taken from meta-analyses or randomized trials. Effectiveness is expressed in terms of YLD averted; indirect benefits such as reduction of lost productivity or lesser pressure on informal caregivers are not taken into account. Whenever assumptions had to be made, we opted for conservative estimates that are unlikely to have resulted in an overly optimistic portrayal of the cost-effectiveness ratios.

Implications for health care provision and use: The model can be used to guide health care system innovation, by identifying those parameters where changes in the uptake, compliance, effectiveness and costs of interventions have the largest impact on the cost-effectiveness of a mental health care system overall. In this sense, the model could assist policymakers during the first stage of decision making on where to make improvements in the health care system, or assist the process of guideline development. However, the improvement candidates need to be assessed during a second-stage 'normative filter', to address considerations other than cost-effectiveness.

INTRODUCTION

Rising health care costs in combination with the economic downturn increase the need for informed policies to improve the cost-effectiveness of the health care systems. Such policies need to reduce the population's disease burden in an economically sustainable way (Lokkerbol et al 2013). Typically, these policies can be directed at changing a number of parameters in the health care system: (1) the uptake of health care interventions, (2) the adherence to interventions, (3) the effectiveness of interventions and (4) the costs of health care interventions. Identifying those parameters that have the greatest leverage for improving the cost-effectiveness of systems overall is a complex task, since there are many diagnostic patient groups and there is a wide choice of interventions available for each group.

Making such complex decisions might be facilitated by using a health economic simulation model. Such a model should be able to combine epidemiological and economic evidence and to conduct 'what-if' analyses in order to generate outcomes that have relevance to policy-making. Indeed, experience with the Australian Assessing Cost-Effectiveness (ACE) models for heart disease, mental disorders and prevention, (Vos et al 2005; Mihalopoulos et al 2011b) and the WHO's CHOICE models (CHOosing Interventions that are Cost-Effective) (Hutubessy et al 2003; Chisholm et al 2004a) indicate that health economic models can contribute to informed policymaking and can play a role in the development of clinical guidelines.

This paper takes a predominantly methodological angle and aims to demonstrate how health economic modelling can help in identifying the most promising directions for a health care system improvement in terms of the overall cost-effectiveness of the health care system. Primary mental health care in the Netherlands is used as a case study for the methodology. Our paper contributes to existing literature by using a health economic model to analyse a system of multiple health care interventions for multiple diagnostic groups, rather than focussing on a single intervention (see for example Smit et al. (2011) and Lokkerbol et al. (2014a)). The disadvantage of evaluating single interventions is that this could lead to sub-optimal solutions rather than the most optimal of the entire health care system. Also, many cost-effective interventions could jointly create a health care system that transgresses budgetary ceilings and put an even larger pressure on the public purse.

First, we will present an overview of the interventions and their corresponding target groups, and focus on the uptake, adherence, cost and effect parameters. The Results section is moulded after the metaphor of 'sowing' and 'reaping'. Some existing health care configurations (combinations of target groups and interventions)

are suboptimal as seen from a cost-effectiveness point of view and one might be tempted to change these in order to improve cost-effectiveness, which requires 'sowing' investments and making targeted changes. Other existing health care configurations are already so cost-effective that it becomes tempting to upscale them to 'reap' more of the benefits offered by these interventions. Where to sow and where to reap is the question and this question is addressed by conducting a series of 'what-if' analyses. The outcome of these analyses is a list of candidates for either up scaling or innovation. Finally, this list of options to improve the health care system should be used as input for a second-stage normative filtering process, where considerations other than cost-effectiveness are taken into account. Our results are of particular interest for the Dutch mental health care system, which is currently facing severe budget cuts and massive reorganisations. For any health care system, our model could serve as a methodological framework for informed health care reforms.

METHODS

Simulation model
We use a health economic simulation model with a one-year time horizon, developed in Excel 2007, to investigate what parameters in the primary mental health care in the Netherlands have the largest impact on the cost-effectiveness of the health care system at national level. Given the time horizon, no discounting was applied. To build the model, we started off with mapping the different types of interventions for all of the different target groups. For each of these "target group – intervention" combinations, we reviewed parameters such as uptake rate, adherence rate, intervention costs and effects, since these four parameters determine the cost-effectiveness. The complete set of target group – intervention combinations with their specific parameter values is called a "health care configuration". The model simulates and quantifies the effect of adjustments of the health care configuration on the overall cost-effectiveness. The section below describes the model in more detail.

Epidemiology
The model discerns six target groups: patients presenting at their GP with symptoms of anxiety, symptoms of stress, symptoms of depression, alcohol dependency/abuse, anxiety disorder and depressive disorder. The model was populated using the GPs'

National Register (LINH), containing data on the diagnostic groups and their treatments and referrals in the Netherlands in the year 2009. Data were based on a representative sample of 40 - 53 well registering GP practices. These data were expansively weighted to reflect the size of the primary mental health care system in the Netherlands, see **table 5.1**. In the model, target groups have to be defined in a mutually exclusive way (to avoid double counting), and can be defined by the user to describe any patient population.

Health care system

Interventions are defined in terms of their uptake rate (percentage of the target group reached), adherence rate (the percentage of patients fully complying with that intervention), intervention costs (the full economic per-patient costs of offering that intervention in euro, €) and effect (the effectiveness of that intervention, expressed as the standardised mean difference d as obtained from meta-analyses or randomized trials). We obtained these parameters for each intervention for each diagnostic target group. To facilitate comparisons across target groups, interventions were described in a generic way, see **table 5.1**.

Uptake rates were extracted from GP registration data as the percentage of patients within each target group receiving a certain intervention. Adherence rates were elicited from a focus group for 'symptoms of depression' and 'depressive disorder'. This focus group consisted of 17 patients with a history of depressive and anxiety disorders who were asked to rate a list of evidence-based interventions in terms of how acceptable they would find each of these interventions. The focus group of patients was very consistent in their judgment (Cronbach's alpha 0.84). For other diagnostic groups, the adherence rates were based on evidence from randomized clinical trials evidence where possible, or else assumed to be 70%, 50%, 30% or 20% for the remaining interventions (see **table 5.1** for details). Costs for each intervention were computed using an auxiliary Costing Tool. Unit cost prices (for GP visits, medication, psychological and e-health intervention) were obtained from the Dutch Guideline for Health Economic Evaluations (Hakkaart-van Roijen et al 2010). The Costing Tool returns the full economic per-participant costs of an intervention and places this estimate within a lower and upper uncertainty bound.

Effects of interventions were extracted from meta-analyses of randomized trials where possible and trials or clinical guidelines otherwise, and expressed in terms of Cohen's standardised effect size d.

Table 5.1 presents the health care configuration corresponding to primary mental health care in the Netherlands, defined in terms of target groups along with their interventions and parameters. This list of interventions forms the basis for the next series of scenario analyses.

Table 5.1: Primary health care system in the Netherlands: intervention costs (in 2009 euro), uptake rate (%), adherence rate (%) and effect (as standardized effect size, d)

	Intervention by target group	Costs	Uptake rate l)	Adherence rate	Effect
T1	**Symptoms of anxiety (N=422,133)**	€	%	%	d
1	Pharmacotherapy	82 a)	63.1	70 [1,2]	1.12 (after [6])
2	Psycho education by GP	69 b)	23.4	70 m)	0.29 (after [7])
3	Psycho education by the GP's assistant	36 c)	0.5	50 n)	0.29 (after [7])
4	Individual psychological intervention	1,154 d)	5.3	70 m)	1.27 (after [6])
5	Psychological group intervention	212 e)	0.0	50 o)	1.27 (after [6])
6	Combination therapy	953 f)	2.4	70 m)	1.34 (after [6])
7	E-health without guidance	175 g)	0.3	70 [3]	0.68 (after [8])
8	E-health with guidance	195 g)	0.0	70 [3]	1.00 (after [1])
T2	**Stress (N=314,750)**	€	%	%	d
1	Pharmacotherapy	82 a)	35.1	70 m)	0.35 x)
2	Psycho education by GP	69 b)	44.3	70 m)	1.02 (after [7])
3	Psycho education by the GP's assistant	36 c)	1.0	50 n)	1.02 (after [7])
4	Individual psychological intervention	1,154 d)	6.6	70 m)	0.515 (after [9])
5	Psychological group intervention	212 e)	0.0	50 o)	0.52 (after [9])
6	Combination therapy	953 f)	2.9	70 m)	0.54 y)
7	E-health without guidance	175 g)	0.0	70 m)	0.32 z)
T3	**Symptoms of depression (N=217,549)**	€	%	%	d
1	Pharmacotherapy	77 h)	43.5	50 [1,2]	0.15 aa)
2	Psycho education by GP	69 b)	36.1	70 m)	0.5 (after [7])
3	Psycho education by the GP's assistant	36 c)	1.5	50 n)	0.5 (after [7])
4	Individual psychological intervention	1,154 d)	6.2	56 p)	0.69 (after [10,11])
5	Psychological group intervention	212 e)	0.0	64 p)	0.31 (after [12,13])
6	Combination therapy	948 f)	4.1	65 p)	0.72 (after [14,15])
7	E-health without guidance	152 g)	3.0	56 p)	0.33 (after [16,17])
8	E-health with guidance	195 g)	0.0	65 q)	0.33 (after [16,17])
T4	**Alcohol dependency / abuse (N=86,916)**	€	%	%	d
1	Pharmacotherapy	86 i)	30.3	50 r)	0.66 (after [18])
2	Psycho education by GP	46 j)	41.7	30 s)	0.32 (after [19])
3	Psycho education by the GP's assistant	36 c)	0.4	20 t)	0.32 (after [19])
4	Individual psychological intervention	1,154 d)	0.3	30 s)	0.46 (after [4,20])
5	Psychological group intervention	212 e)	0.0	20 u)	0.28 (after [21])
6	Combination therapy	960 f)	1.9	30 s)	0.21 (after [22])
7	E-health without guidance	152 g)	15.0	30 s)	0.31 (after [23])
8	E-health with guidance	764 k)	0.0	30 v)	0.585 (after [5])
T5	**Anxiety disorder (N=214,609)**	€	%	%	d
1	Pharmacotherapy	78 h)	62.8	70 [1,2]	1.12 (after [6])
2	Psycho education by GP	46 j)	22.1	70 m)	0.29 (after [7])
3	Psycho education by the GP's assistant	36 c)	0.5	50 n)	0.29 (after [7])
4	Individual psychological intervention	1,154 d)	4.5	70 m)	1.27 (after [6])
5	Psychological group intervention	212 e)	0.0	50 o)	1.27 (after [6])
6	Combination therapy	949 f)	4.1	70 m)	1.34 (after [6])
7	E-health without guidance	152 g)	0.6	70 [3]	0.68 (after [8])
8	E-health with guidance	195 g)	0.0	70 [3]	1.00 (after [1])

T6 Depressive disorder (N=535,582)	€	%	%	d
1 Pharmacotherapy	*76 h)*	*75.2*	*50 [1,2]*	*0.72 (after [14,15,24])*
2 Psycho education by GP	*46 j)*	*12.5*	*70 m)*	*0.5 (after [7])*
3 Psycho education by the GP's assistant	*36 c)*	*0.2*	*50 n)*	*0.5 (after [7])*
4 Individual psychological intervention	*1,154 d)*	*2.8*	*68 p)*	*0.69 (after [10,11])*
5 Psychological group intervention	*212 e)*	*0.0*	*50 p)*	*0.69 bb)*
6 Combination therapy	*947 f)*	*3.3*	*65 p)*	*1.05 (after [25])*
7 E-health without guidance	*152 g)*	*1.2*	*22 w)*	*0.32 (after [16,17])*
8 E-health with guidance	*195 g)*	*0.0*	*43 p)*	*0.32 (after [16,17])*

a) 3-4 contacts with GP of 10 min each plus 12 months pharmacotherapy at 6.75 per month
b) 3 contacts with GP of 10 min each
c) 1 contact with GP's assistant of 30 min
d) 1 referral by GP, 7 (6-8) consults with psychologist of 45 min each
e) 1 referral by GP, 7 (6-8) consults with psychologist in group of 6 (4-8) persons
f) 1 referral by GP, 5 consults with psychiatrist of 20 min each, 12 months pharmacotherapy
g) Based on tariffs from a Dutch e-health provider (www.mentalshare.com)
h) 3-4 contacts with GP of 10 min each plus 12 months pharmacotherapy at 4.42 per month
i) 3-4 contacts with GP of 10 min each plus 12 months pharmacotherapy at 7.25 per month
j) 2 contacts with GP of 10 min each for screening, personalized feedback and possibly referral
k) 1 referral by GP, 5 sessions with online therapist, website costs
l) Based on register data by GPs
m) Unknown, assumed 70%
n) Unknown, assumed 20% lower than compliance GP
o) Unknown, assumed 20% lower than compliance individual psychological intervention
p) Elicited from a panel of health care users
q) Unknown, assumed somewhat higher than compliance e-health unguided
r) Estimation after results for acamprosate and naltrexone in (GGZ guidelines 2009)
s) Unknown, due to generally low compliance in patients with alcohol disorder assumed to be 30%
t) Unknown, assumed to be 10% lower than compliance GP
u) Unknown, assumed to be 10% lower than compliance individual psychological intervention
v) After Blankers et al. (2011)
w) Unknown, assumed half the compliance of e-health guided
x) Assumed arbitrarily low effect size of 0.35.
y) Assumed somewhat higher than psychotherapy
z) Conservatively estimated based on effects relating e-health unguided in other target groups
aa) Conservatively estimated a very low effect size 0.15
bb) Assumed equal to the effect of psychological group intervention

1) GGZ guidelines 2011a
2) GGZ guidelines 2011b
3) Haverman et al 2009
4) GGZ guidelines 2009
5) Blankers et al 2011
6) Issikadis et al 2004
7) Bakker et al 2007
8) Meulenbeek et al 2010
9) Van der Klink et al 2001
10) Cuijpers et al 2007
11) Cuijpers et al 2008c
12) Cuijpers 1998
13) Cuijpers et al 2008a
14) Fournier et al 2010
15) Kirsch et al 2008
16) Spek et al 2007a
17) Cuijpers et al 2008b
18) Kosten and Connor 2003
19) Beich et al 2003
20) Walters 2000
21) Tonigan et al 1996
22) Mann 2004
23) Riper et al 2008
24) Arroll et al 2005
25) Cuijpers et al 2009

Assessing health gains

Health care interventions aim to generate health gains in the population, or equivalently, to reduce disease burden. At population level, averting disease burden can be captured in years lived with disability (YLD) averted. Using this perspective, benefits outside the health care system, such as productivity gains or a reduced demand for informal care, are not taken into account. To quantify YLD averted, the health economic simulation model makes use of the standardized mean difference (Cohen's d), because this statistic is widely available in many meta-analyses of randomized trials of treatments. One could think of this effect size as a reduction in symptom severity, expressed in standard units, d. A shift in symptom severity (of size d) needs to be translated into a corresponding shift in the disability weight, DW, to calculate treatment impacts on YLD averted. For a range of mental disorders Sanderson et al. (2004) estimated "conversion factors" that translate a shift of size d into a corresponding shift in the DW. Because our model contains multiple types of disorders, we use a conversion factor of 0.11, with a standard error of 0.0105 to explicitly take uncertainty into account, based on a conservative estimation across different conditions. The change in DW is then multiplied by the appropriate number of person years (number of people multiplied by their time spent in the condition) to arrive at an estimate of the number of YLD averted by the intervention.

Thus, total YLD averted are calculated by multiplying group size by uptake rate, adherence rate, effect size d and the conversion factor, for each target group intervention combination.

Assessing costs

Offering health care interventions in order to avert YLD in the population comes at a cost. Since we look at direct medical intervention cost, this equals the full economic cost price of the required resources, where resource use is guideline concordant. (GGZ guidelines 2009; 2011a; 2011b) In the model, total costs are calculated as the cost per intervention multiplied by the uptake rate and the size of the target group, for each target group - intervention combination.

Cost-effectiveness analysis

The most commonly used measure in health economics to compare two scenarios (a base case scenario subscripted 0 and an alternative scenario subscripted 1) in terms of cost-effectiveness, is the incremental cost-effectiveness ratio (ICER):

ICER = (C_1 - C_0) / (E_1 - E_0).

For both scenarios the health economic model runs 1,000 simulations, taking uncertainty into account around cost and effect parameters. After the simulations, the ICER is used as a measure to summarize the health economic impact of changing from scenario 0 to scenario 1. The difference in costs for both scenarios (C_1-C_0) is divided by the difference in effect for both scenarios (E_1-E_0). In case of investigating an additional intervention in the alternative scenario compared to the base case scenario, the ICER can therefore be interpreted as the additional cost (C_1-C_0) per additional unit effect (E_1-E_0). Policymakers can use the ICER to decide whether the additional effects of a new intervention are worth the additional costs.

However, we compare the health care configuration (**table 5.1**) with exactly the same configuration with just a single parameter (uptake rate, adherence rate, costs or effect) changed. Under these conditions the ICER is inappropriate, because the difference between two scenarios generally amounts to only a difference in cost (when only one cost parameter is changed) or only a difference in effect (when only one adherence rate or effect parameter is changed). Except for the situation where an uptake parameter is changed, there is a change in only costs or effects, which leads to ICER values of either 0 (when costs do not change) or infinity (when effects do not change between two scenarios). To still be able to make comparisons, we do not use the current health care configuration (**table 5.1**) as the base case scenario, but rather the scenario of "no care", which entails zero intervention costs and zero effects. With this specific base case scenario, the ICER simplifies to:

ICER = cost / effect,

where cost and effect refer to the simulated costs and effects of a specific health care configuration. In our analysis, we thus compare the ICER of current care versus no care, with the ICERs of current care with one parameter improved versus no care.

For each specific health care configuration, the simulation model generates vectors of costs and YLD averted, taking uncertainty around cost and effect parameters into account. Uncertainty in the model is a key issue. After all, uncertainty regarding the value of each of the input parameters may carry over into the results and needs to be an explicit part of the output. Therefore, the model allows parameters to change freely over a range of likely values. The model takes uncertainty around cost and effect parameters into account by running 1,000 simulations, where the costs and effects of each intervention are randomly drawn from an underlying distribution (a gamma distribution for the costs and a normal distribution for the standardised effect sizes d). The outcomes of each simulation (costs and YLD averted for each scenario)

are computed and stored in the model's memory. In a final step, all simulated outcomes are evaluated simultaneously. At that point, the modelled outcomes are no longer deterministic, but probabilistic and capture the uncertainty in the cost-effectiveness estimates due to uncertainty in the input parameters.

By dividing costs by health gains for both scenarios, the cost-effectiveness of both scenarios can be compared, making it possible to analyse the effect of one specific parameter adjustment on the overall cost-effectiveness – thus capturing the impact of each of the parameters on the cost-effectiveness of the health care system. In this way the model can be seen as an algorithm for finding directions toward sustainable system innovation.

Sensitivity analysis

The algorithm used for obtaining directions on where to invest and where to reap is basically a form of sensitivity analysis. The model changes one single parameter at a time and recalculates the overall cost-effectiveness by running 1,000 simulations. After running the model for every possible parameter adjustment, corresponding health care interventions of the parameters with the largest impact on the overall cost-effectiveness can be selected. This is done for each of the parameter types: uptake, adherence, cost and effect.

RESULTS

For each of the four types of parameters (uptake, adherence, cost and effect), it is shown in what target group – intervention combination a 5% parameter change generates the largest impact on the health care system's cost-effectiveness.

Uptake

The model changed the uptake rate of each distinct intervention by 5% (e.g. from 20% to 25%). The five interventions that resulted in the largest improvement of the cost-effectiveness are shown in **table 5.2**.

Table 5.2: Top-5 target group – intervention combinations that reduce the overall cost/effect ratio the most when uptake rate is increased by 5%

	Target group - intervention	Improvement cost / effect ratio
1	Stress – Psycho education by GP	1.18%
2	Symptoms of anxiety - Pharmacotherapy	1.11%
3	Symptoms of anxiety – Psychological group intervention	1.04%
4	Stress – Psycho education by GP's assistant	0.91%
5	Symptoms of anxiety – E-health with support	0.74%

These interventions provide the greatest leverage for improving the cost-effectiveness of the health care system overall when their uptake rate is changed, where increasing the uptake rate of psycho-education by the GP for patients presenting with stress seems to have the biggest impact. Since these interventions are most efficient in increasing health, these interventions can be seen as the best candidates to reap the benefits of the existing health care configuration.

Adherence

The model was used to recalculate the cost-effectiveness of the health care system after increasing the adherence rate of each distinct intervention by 5%. The five interventions that resulted in the largest efficiency gains are presented in **table 5.3**.

Table 5.3: Top-5 target group – intervention combinations that reduce the overall cost/effect ratio the most when adherence rate is increased by 5%

	Target group - intervention	Improvement cost / effect ratio
1	Symptoms of anxiety – Pharmacotherapy	2.17%
2	Depressive disorder – Pharmacotherapy	1.27%
3	Anxiety disorder – Pharmacotherapy	1.11%
4	Stress – Psycho education by GP	1.04%
5	Stress – Pharmacotherapy	0.29%

These are the interventions that generate the largest improvements in the cost-effectiveness of the health care system when their adherence rates are increased, where increasing the adherence of pharmacotherapy for patients presenting with symptoms of anxiety seems to have the biggest impact. It is worth noting that these results are largely brought about by the size of the target group, and the uptake rates and effect sizes of these interventions, which act as a leverage for adherence. These interventions can therefore be seen as the best candidates within the health care system that could benefit most from innovation aimed at improving the patient's adherence, for example monitoring and promoting compliance with pharmacy use with help of smart packages and mobile technologies that alert a patient when medication needs to be taken.

Costs

The model was also used to recalculate the cost-effectiveness of the health care system, after decreasing the costs of each intervention by 5%. The five interventions that resulted in the most substantial improvement of the cost-effectiveness are presented in **table 5.4**.

Table 5.4: Top-5 target group – intervention combinations that reduce the overall cost/effect ratio the most when cost is decreased by 5%

	Target group - intervention	Improvement cost / effect ratio
1	Depressive disorder – Pharmacotherapy	1.06%
2	Symptoms of anxiety – Pharmacotherapy	0.83%
3	Anxiety disorder – Pharmacotherapy	0.36%
4	Stress – Pharmacotherapy	0.35%
5	Symptoms of depression – Pharmacotherapy	0.25%

The interventions that increase the cost-effectiveness of the health care system most when decreasing cost can be interpreted as the interventions requiring the most money to offer. Reducing the costs of these interventions is therefore associated with the largest improvement of the cost-effectiveness of the health care system, where decreasing the cost of pharmacotherapy for patients presenting with depressive disorder seems to have the biggest impact. Cost reductions could be achieved, for example, by more often prescribing low cost medication (generics rather than brands, or first generation rather than second generation antidepressants), as long as the effectiveness is not compromised.

Effect

Finally, the model was used to recalculate the cost-effectiveness of the health care system, after increasing the effect of each intervention by 5%. The five interventions that resulted in the largest improvement of the cost-effectiveness are shown in **table 5.5**.

Table 5.5: Top-5 target group – intervention combinations that reduce the overall cost/effect ratio the most when effect is increased by 5%

	Target group - intervention	Improvement cost / effect ratio
1	Symptoms of anxiety – Pharmacotherapy	1.53%
2	Anxiety disorder – Pharmacotherapy	0.80%
3	Stress – Psycho education by GP	0.73%
4	Depressive disorder – Pharmacotherapy	0.64%
5	Stress – Pharmacotherapy	0.20%

These are the interventions that increase the cost-effectiveness of the overall health care system most when increasing the clinical effectiveness, for example via careful titration of the optimal doses for individual patients. Increasing the effectiveness of pharmacotherapy for patients presenting with symptoms of anxiety seems to have the biggest impact on the cost-effectiveness of the health care system.

CONCLUSION

Main findings

The main finding of this study is that a health economic simulation model can be used to systematically scan a health care system in order to identify parameters that help to improve the cost-effectiveness of the health care system overall. Increasing uptake can be seen as reaping the benefits of interventions that are currently cost-effective. Improving adherence, costs or effectiveness can be seen as investing in those areas that potentially have the biggest leverage on the health care system as a whole.

With regard to the Dutch primary mental health care system, we found that the cost-effectiveness will benefit most from 1) increasing the uptake of psycho-education by a GP for people presenting with stress, 2) increasing the adherence to pharmacotherapy in people presenting with symptoms of anxiety, 3) decreasing the cost of pharmacotherapy for people presenting with depressive disorder, and 4) increasing the effectiveness of pharmacotherapy for people presenting with symptoms of anxiety. These target group – intervention combinations are the most promising in their own class of parameters, which does not mean that these are the four target group – intervention combinations with the most potential overall to improve the cost-effectiveness of the health care system. The shortlists of options could be used as input for policymakers or to support the guideline development process. Input can be assessed in a second-stage 'normative' (medical-ethics) filter, taking considerations other than cost-effectiveness into account, such as equity, acceptability, appropriateness, feasibility and strength of evidence.

Strengths and limitations

One of the strengths of this paper is that a simulation model was used to help organize vast fields of knowledge across the disciplines of epidemiology, medicine, psychology and health economics. Evidence on effect parameters were taken from randomized clinical trials, meta-analyses, and evidence-based clinical guidelines. In addition, a model makes all necessary information available in a dynamic form, which

makes it possible to conduct 'what – if' analyses. This could be of assistance to policy formulation and setting research agendas for health care innovation especially when the subject matter is intrinsically complex.

An additional strength is that our model is based on Dutch General Practitioners' Register (LINH) data. For the year 2009, registration data on patients presenting with symptoms of anxiety, stress, symptoms of depression, alcohol dependence and abuse, anxiety disorders and depressive disorder were used, along with data on the treatments they received and referrals to specialist care.

Our simulation model may have specific merits for setting research agendas. It helps to identify those parameters that have the largest impact on population health and costs. Parameters with the largest effect on cost-effectiveness are obvious candidates for innovation, but when some of these parameters are surrounded by much uncertainty, then it is recommended to conduct empirical research with the aim to reduce uncertainty in the parameter estimates.

A limitation of our findings is that our results might be biased towards the more commonly used interventions such as pharmacotherapy, since these have substantial impact on the cost-effectiveness of the health care system as a whole by force of the large number of patients receiving such interventions. After all, both the size of a target group and the uptake rate of an intervention largely determine the importance of the interventions with respect to the current cost-effectiveness. In general, our results should therefore not be interpreted as an identification of the interventions with the greatest potential of being cost-effective in and by themselves, but as the interventions with the greatest potential to leverage the cost-effectiveness at the macro level of a specific health care system. This makes our algorithm valid to policymakers at national level. Still, our algorithm does tend to overemphasize the more commonly used interventions at the expense of potentially interesting, less commonly used interventions, such as e-health interventions, and the smaller target groups, such as patients with alcohol dependency/abuse. This 'bias' reflects the epidemiology of the Dutch population and the historical process that shaped the Dutch health care system. Because of such historic idiosyncrasies, each country will have its own unique health care configuration. Since the outcomes of our algorithm are conditioned on this configuration, the results are likely to be different for different health care systems. Consequently, our findings cannot be generalised to other health care systems in other countries. For other health care systems, be it at national or local level, our algorithm should be individually applied.

Another limitation of our model is the availability of data. As with any modelling work, the quality of the output depends on the quality of the input. Although the main purpose of our paper is to demonstrate the methodology, it is

important to realize the limitations regarding the data used. We based the size of target groups and uptake rates on the Dutch GP Register data. Data were based on a representative sample of 40 - 53 well registering GP practices. Selecting only well registering GP practices is believed to contribute to a representative sample. At the same time, there is a risk that numbers are underestimated, because some complaints and disorders (especially the milder ones) are not always recognised and in addition, when recognised not always registered. Such downward biases, when not equally distributed across every diagnostic group, can alter the findings of the study, especially since findings depend on the relative sizes of the target groups. Adherence rates and some of the effect parameters are not supported by robust evidence. As an alternative, we elicited adherence rates from a single focus group for patients with symptoms of depression and depressive disorders to arrive at an estimate for adherence rates. Even though the focus group was consistent in their judgment, a series of structured interviews with patients could yield more reliable estimates. For the other target groups we had to rely on mere assumptions, rendering our results regarding adherence rates tentative.

The comparability of effect sizes between different interventions is another concern. Since our evidence was taken from meta-analyses, RCTs and clinical guidelines, effect parameters reflect the effect of an intervention relative to different comparator conditions, such as placebo, waiting list, or care as usual, where usual care may differ between countries. Limiting the evidence-base to countries that are more or less similar to the Dutch health care system may help, as well as obtaining the effect sizes indirectly from a comparator-adjusted network meta-analysis to better handle differences in the comparator condition.

The use of an average conversion factor across different target groups rather than target group specific conversion factors could introduce a bias towards recommending interventions from target groups with below-average conversion factors. Since specific conversion factors are not known for every target group, estimating and using target group specific conversion factors is believed to introduce a new bias. Furthermore, with each simulation a conversion factor is randomly drawn from a normal distribution, explicitly taking into account the uncertainty around this parameter.

Another limitation of our model is that by assessing the cost-effectiveness of the health care system after improving one single parameter, it is implicitly assumed that all other parameters are independent and thus remain the same. In reality, parameters could be correlated. For example, improved adherence might lead to lower costs, as a better adherence could lead to less required resources in order to achieve the required result. However, since we only look at minor adjustments, (5%

changes) in the health care system, we do not expect this limitation to have a substantial impact on our outcomes.

Another limitation of our model is its narrow focus on health economic implications. First of all, we do not look at costs outside the health care system, such as work productivity or informal care, which could have a significant impact (Davidson 2009). Also, health care costs other than intervention costs are not taken into account. Next to that, considerations such as equity and medical ethics are not taken into account. Therefore, our results should be seen as input for a second-stage normative filtering process, as in Mihalopoulos et al. (2011b), where considerations other than cost-effectiveness can be taken into account. At this point, it should also be questioned how realistic it is to further improve relatively large effect sizes, or to further decrease relatively cheap interventions.

A final limitation is that the model only offers directions for health care system optimization. Once a parameter has been identified as having a great leverage on the cost-effectiveness of the health care system, then one still needs to devise practical strategies that help produce the intended efficiency gains in the clinical setting.

Implications

In times of economic downturn and budget cuts in health care, it is important for policymakers in the health care sector to design strategies that are likely to have the most beneficial impact on the cost-effectiveness of the health care system as a whole. Our health economic model helps with the identification of where money is best spent in order to reap the benefits of current interventions, and to direct innovations such that the largest cost-effectiveness improvements in health care can be sowed. The methodology could be used to support the process of multidisciplinary guideline development, where our results can be considered taking into account other perspectives such as equity, appropriateness, acceptability, feasibility and strength of evidence by different agencies involved, such as health care professionals and patients. At this stage we are reluctant to make recommendations for policymakers, but begin to see a methodology that eventually may help to identify opportunities for increasing the cost-effectiveness of health care systems.

Improving the cost-effectiveness of health care by simultaneously investing and disinvesting in interventions

Authors: Lokkerbol J, Weehuizen R, Cuijpers P, Mihalopoulos C, Smit F

Abstract

Objective: In the context of the economic downturn and budgetary constraints it is important to develop an algorithm for identifying opportunities to create a more cost-effective health care system.

Methods: A health economic substitution algorithm was developed to assess the potential of pairs of interventions within the current intervention mix to jointly create a more cost-effective health care system by (partly) substituting one intervention by an alternative intervention. This substitution algorithm is applied to the Dutch health care system for patients stratified by mild, moderate and severe depressive disorder for illustrative purposes.

Results: The algorithm identified 11 intervention pairs within the current intervention mix of the health care system for major depression in the Netherlands with the potential to arrive at a more cost-effective health care system. For each intervention pair it is explicated how substitution of the interventions can improve the overall costs and/or health effects of the health care system.

Conclusions: The algorithm can be used to guide health care system improvement, by providing policymakers with a list of propositions to obtain a more cost-effective health care system. Making substitutions within the existing health care package could potentially pose a smaller implementation barrier than improving the cost-effectiveness of the health care system by adding new interventions. Nevertheless, once options for improving the cost-effectiveness of the health system have been identified, a 'second-stage normative filtering' process needs to be applied to take into account other considerations such as medical-ethics, feasibility, interchangeability, acceptability, and appropriateness of the propositions.

INTRODUCTION

Rising health care costs (OECD stats, accessed 05-02-2014) in combination with the economic downturn increases the need for policymakers to make informed decisions with both population health and macro-economic implications in mind. In order to manage disease burden (Lokkerbol et al 2013) in a sustainable manner, resources should ideally be allocated in such a way as to make the health care system cost-effective in an optimal way.

Finding ways to improve the cost-effectiveness of the existing health care system using health economic modelling is not new. Health economic models combining epidemiological and economic evidence help to assess cost-effectiveness and facilitate complex decision-making from the policy perspective. Experience with the Australian Assessing Cost-Effectiveness (ACE) models for heart disease, mental disorders and prevention (Vos et al 2005; Mihalopoulos et al 2011b), and the WHO's CHOICE models (CHOosing Interventions that are Cost-Effective) (Hutubessy et al 2003; Chisholm et al 2004a) indicate that health economic models may have value for such purposes.

Traditionally, improving cost-effectiveness has a focus on adding new and presumably more cost-effective interventions to care as usual (for example Smit et al 2006b; 2011; Valmaggia et al 2009; van Spijker et al 2012; Lokkerbol et al 2014a). However, adding new interventions requires additional investment and increases pressure on budgets. This trend towards expanding the health care system goes hand in hand with a trend of increasing health care expenditure. Next to pressuring budgets, there are two important downsides to this route to innovation. One is that new interventions can pose a challenge in terms of implementation, as new interventions may require new infrastructures, additional training, and often introduce uncertainty around the actual benefits of the intervention outside the experimental setting due to the possible gap between efficacy and effectiveness (Proctor et al 2009). Another drawback is that the cost-effectiveness of new interventions is often compared with care as usual. Care as usual, however, can be sub-optimal in terms of cost-effectiveness (Andrews et al 2004; Vos et al 2005). Comparing a new intervention to care as usual can therefore exaggerate the cost-effectiveness of this new intervention, thereby reducing the practical value of such comparisons. In this context, it is important to first optimize care as usual by addressing the question whether the existing package of interventions for a diagnostic target group is optimal from a health economic point of view, before looking at implementing new interventions.

Optimizing the current intervention mix could be done using disinvestment. Although disinvestment poses difficult scientific, political and ethical challenges (Pearson and Littlejohns 2007), disinvestment is growing as an international priority, in order to improve quality of care and its sustainability (Elshaug et al 2008), or as a means to alleviate financial pressure on health services (Garner et al 2013).

We present a substitution algorithm that identifies intervention pairs in the current intervention mix of a health care system that can improve the overall costs and/or effects of the health care system via health care substitution. We build on the approach outlined in Sendi et al. (2002), who propose a graphical framework (which they call the 'decision making plane', similar to the planes presented in **figure 6.1**), to investigate how investment options can be financed by downscaling existing programs, resulting in overall improved costs and/or effects. In this paper, such a health care configuration with improved costs and/or effects is referred to as a 'dominant health care configuration'. Our algorithm contributes by looking at the theoretically 'optimal' degree of substitution, as well as presenting a case study to demonstrate the algorithm. Only interventions competing for the same diagnostic target group are compared, such that substitution is a viable option from a clinical perspective. Opportunities, once identified, can serve as input for a second-stage normative filtering process to also consider issues other than cost-effectiveness, such as medical ethics considerations.

The first part of the methods section describes the substitution algorithm used to identify the intervention pairs that have the potential to constitute a dominant health care configuration, where the health care configuration refers to the specific mix of interventions that constitute the health care system. The second part of the methods section briefly describes the health economic model used to demonstrate the algorithm, as this model has been described elsewhere in detail (Lokkerbol et al 2014a). The results section presents the intervention pairs with their potential for improving the cost-effectiveness of the health care system for major depression in the Netherlands. The last section discusses strengths and limitations of the method used, as well as the broader implications of investing and disinvesting in a health care system. Our results are particularly relevant for the Dutch mental health care system, which is currently facing budget constraints and large-scale reorganisations. Nevertheless, the principles used and lessons learned in the Dutch context may have value for other health care systems in other countries.

METHODS

The algorithm constructs a list of intervention pairs for each diagnostic target group along with their improvement potential by:

- systematically scanning every possible combination of two interventions in the health care system for the same diagnostically homogeneous target group,
- assessing each intervention pair in terms of the potential increase in net benefit as compared to care as usual it can achieve via substitution, and
- ranking this list from highest to lowest additional net benefit.

Every two interventions with a different cost-effectiveness ratio (CE-ratio) have the potential to create a dominant health care configuration by scaling up the intervention with the lower CE-ratio while scaling down the intervention with the higher CE-ratio. In the (common) situation where the health care system has no interventions with identical CE-ratios, every single intervention pair has the potential to create a dominant health care configuration.

First, each pair is assessed to determine which intervention should be invested in (lower CE-ratio) and which intervention should be disinvested (higher CE-ratio). Second, care as usual is compared to the health care system after substitution of the intervention pair.

Since there are many degrees in which two interventions can be substituted, we need to define the 'optimal substitution' to be able to compare different intervention pairs in their ability to improve the health care system. We define this theoretical optimum as the substitution associated with the greatest net benefit (total health gains multiplied by the willingness-to-pay (WTP), representing the value society attaches to one health unit, minus total costs). When the maximum net benefit of each intervention pair is assessed, intervention pairs can be ranked in order to present the intervention pairs with the greatest potential to increase the net benefit of the health care system.

Substitution algorithm

Let us assume we construct a dominant health care configuration by scaling up intervention 1 and simultaneously scaling down intervention 2. This leads to a dominant health care configuration when the health effects gained by scaling up intervention 1 are more than (at least as much as) the health effects lost when scaling

down intervention 2, and when the budget freed up by scaling down intervention 2 is at least as much as (more than) the budget needed to scale up intervention 1.

To identify the degree of substitution (full or partial) that will lead to the largest improvement in overall cost-effectiveness, we start off with mapping the additional costs and health effects, defined as C_1 and E_1 respectively, when scaling up intervention 1, and the budget freed up and health effects lost, defined as C_2 and E_2 respectively, when scaling down intervention 2.

Finding the theoretical optimum for each intervention pair is not straightforward, as it depends on the CE-ratio of both interventions, the magnitude of C_1, E_1, C_2, E_2 and WTP. Furthermore, we explicitly pose the restriction that the new health care system has to have at least as much health effects, while spending the same or less budget, as we are interested in the potential of the health care system to achieve such a situation.

Only pairs where C_1/E_1 is below the WTP threshold are considered, since investing in these interventions will lead to health gains at a cost society is willing to pay for. Imposing this restriction prevents the algorithm from proposing improvements based on scaling up interventions that are considered 'bad value for money', although this would be at the expense of interventions that offer even worse value for money.

Furthermore, improving the cost-effectiveness of the health care system by investing in intervention 1 while disinvesting in intervention 2 to construct a dominant health care configuration implies that $C_1/E_1 < C_2/E_2$.
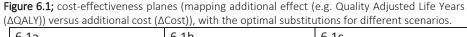
When considering the possible substitutions for each intervention pair, we put a cap on the extent to which intervention 1 can be scaled up, equal to the coverage rate of intervention 2. For reasons of generalizability, we refer to scaling up intervention 1 with the coverage rate of intervention 2 as 'full investment'. We impose this restriction because we are primarily interested in how substitution could lead to a dominant health care configuration. If we would allow for a larger increase in the coverage rate of intervention 1, we are no longer strictly looking at substitution but also to expanding the reach of the health care system. As will be shown in the results section, sometimes an increase in the reach is needed to ensure that the health care system does not worsen in terms of health effects.
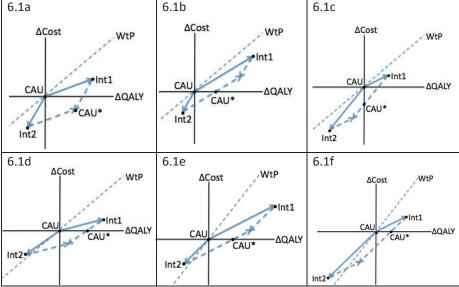
Under the restriction that our theoretical optimum needs to have more (at least as much) health effects for a similar (smaller) budget, we discern six situations:

- The most intuitive situation is where intervention 2 achieves health effects at a cost that is higher than the WTP threshold. In this case, health effects as obtained by intervention 2 are considered 'too expensive'. Besides wanting to invest in intervention 1, we would like to disinvest in intervention 2. Whether a full investment and full disinvestment is feasible depends on whether the budget freed up by fully disinvesting intervention 2 is sufficient to cover the additional cost of fully investing in intervention 1, and whether the additional health effects as obtained by fully investing in intervention 1 are more than the health effects lost due to fully disinvesting intervention 2. We discern three possible scenarios:

    1. $C_1 < C_2$ and $E_1 > E_2$: the theoretical optimum lies in the 'corner solution' of full investment in intervention 1 and full disinvestment of intervention 2. In **figure 6.1a** it can be seen that full investment in intervention 1 moves care as usual (CAU) to int1 in the upper right quadrant, as investing in intervention 1 increases both health and costs. Simultaneous disinvestment in intervention 2 shifts the health care system down to the bottom right quadrant (from int1 to CAU*), resulting in a health care system that is cheaper than the original CAU, with larger health gains. This dominance is achieved because disinvesting in intervention 2 frees up more budget than required for investing in intervention 1, while investing in intervention 1 results in health gains that more than compensate for the decrease in health due to disinvesting in intervention 2.

    2. $C_1 > C_2$ and $E_1 > E_2$: investing in intervention 1 requires more budget than disinvesting in intervention 2 can provide for. The health care system with the largest net benefit is obtained by fully disinvesting intervention 2, while investing in intervention 1 up to the point where the overall change is budgetary neutral, see **figure 6.1b**.

    3. $C_1 < C_2$ and $E_1 < E_2$: not enough health effects are gained when investing in intervention 1 to offset the health effects lost when fully disinvesting intervention 2. The health care system with the largest net benefit is obtained by fully investing in intervention 1, and disinvesting in intervention 2 up to the point where the effect on health is neutral, see **figure 6.1c**.

The last scenario, with $C_1 > C_2$ and $E_1 < E_2$, is not considered, as this situation implies that intervention 2 should be scaled up rather than scaled down.

- In the other situation intervention 2 achieves health effects at a cost that is lower than the WTP threshold. In this case, health effects as obtained by intervention 2 are considered 'good value for money'. Even though intervention 2 is considered to offer good value for money, disinvestment in intervention 2 is desirable from a relative point of view (since we imposed $C_1/E_1 < C_2/E_2$). Again, there are three possible scenarios:

4. $C_1 < C_2$ and $E_1 > E_2$: the theoretical optimum lies in full investment in intervention 1 while disinvesting in intervention 2, up to the point where the overall change is budgetary neutral, see **figure 6.1d**.

5. $C_1 > C_2$ and $E_1 > E_2$: investing in intervention 1 requires more budget than disinvesting in intervention 2 can provide for. Therefore, the theoretical optimum lies in fully disinvesting intervention 2, while investing in intervention 1 up to the point where the overall change is budgetary neutral, see **figure 6.1e**. This scenario is qualitatively equal to scenario 2.

6. $C_1 < C_2$ and $E_1 < E_2$: not enough health effects are gained when investing in intervention 1 to offset the health effects lost when fully disinvesting intervention 2. The theoretical optimum lies in fully investing in intervention 1, and disinvesting in intervention 2 up to the point where the change is budgetary neutral, see **figure 6.1f**.

Figure 6.1; cost-effectiveness planes (mapping additional effect (e.g. Quality Adjusted Life Years (ΔQALY)) versus additional cost (ΔCost)), with the optimal substitutions for different scenarios.

CAU = Care as usual; WTP = willingness-to-pay indifference curve (in terms of cost-effectiveness, alternative health care systems on this curve are considered 'as good as' CAU, health care systems above the curve are 'worse than' CAU, health care systems below the curve are considered 'better than' CAU); Int1 is the health care system associated with CAU after investing in intervention 1; Int 2 is the health care system associated with CAU after disinvesting in intervention 2; CAU* represents the health care system with the highest net benefit when substituting intervention 2 for intervention 1, given the restrictions.

It is worth noting that each intervention pair constructs a parallelogram containing all possible health care systems that can be constructed by combining interventions 1 and 2. The theoretical optimum is related to the willingness-to-pay indifference line, which represents all the health care systems that are considered 'as good as' care as usual, since the change in health effects is obtained by the exact cost that society values these health effects. The theoretical optimum has the largest orthogonal distance to this indifference line.

After locating each theoretical optimum, each optimum is valued in terms of the increase in net benefit compared to care as usual: Increase in net benefit = decrease in cost + increase in health * WTP.

Our approach requires that interventions need to be defined in such a way that each patient is allocated to one single intervention category (see table 6.1 for examples of this). This condition contributes to making the interventions mutually exclusive and

independent, which makes the interventions freely substitutable, and indeed lets a combination of two interventions create the area of potential health care systems as depicted by the parallelograms in **figure 6.1**.

Once each intervention pair is defined in terms of one of these six categories and valued accordingly in terms of additional net benefit, all pairs can be ranked in terms of their theoretical improvement potential, thereby providing policymakers with propositions for the innovation agenda. As mentioned earlier, these propositions for improving the cost-effectiveness need to pass a 'second-stage normative filter' to take considerations other than cost-effectiveness into account (Mihalopoulos et al 2011b).

Major depression in the Netherlands
Major depression in the Netherlands is used as a case study to demonstrate the substitution algorithm presented in this paper. The starting point is a health economic simulation model describing care as usual for patients aged 18-65 diagnosed with major depression in the Netherlands. The model is explained briefly here, and is described in detail in Lokkerbol et al. (2014a).

A Markov-model was developed for diagnostic target groups stratified by depression severity (mild, moderate, severe). Interventions for each target group, such as psychological interventions or pharmacotherapy, were defined using the pertinent clinical guidelines (GGZ guidelines 2011b) as a starting point. The simulation model combines:
- the epidemiology of depression (incidence, prevalence, recovery, recurrence and excess mortality, based on Dutch cohort data (Bijl et al 2002; de Graaf et al 2012b),
- clinical evidence of treatment effectiveness (based on meta-analyses or randomized trials for each of the interventions, for example Arroll et al 2005; Vittengl et al 2007; Cuijpers et al 2007; 2008c; 2009; Ekers et al 2008; Kirsch et al 2008; Fournier et al 2010),
- data on coverage rates (elicited from a health care provider focus group),
- data on treatment adherence (elicited from a patient focus group or based on expert opinion), and
- the full economic per-patient costs of offering interventions (based on resource use and standard Dutch unit cost prices (Hakkaart-van Roijen et al 2010)).

In this way, the health economic model describes care as usual, as far as it is evidence-based, deemed appropriate by health care providers and acceptable by patients. Besides defining usual care, the model also allows for defining alternative health care systems, thus making it possible to explore the impact on costs and effects during a five-year time horizon when scaling up or scaling down interventions by changing their coverage rates in the alternative scenario.

Assessing intervention pairs

The health care system for the diagnostic groups mild, moderate and severe depression, are defined in Lokkerbol et al. (2014a) in terms of 3, 3 and 4 generic interventions respectively (see **table 6.1**).

Table 6.1: Generic interventions for major depression (mild, moderate, severe) in the Netherlands

	MILD DEPRESSION	
Mild-1	E-health intervention (supported), 8 sessions	
Mild-2	Online psychotherapy, 10 sessions CBT	
Mild-3	Individual psychotherapy, primary care, 8 sessions	
	MODERATE DEPRESSION	
Mod-1	E-health intervention (supported), 8 sessions	
Mod-2	Online psychotherapy, 14 sessions CBT	
Mod-3	Individual psychotherapy, primary care, 8 sessions	
	SEVERE DEPRESSION	
Sev-1	Individual psychotherapy, outpatient care, 8-24 sessions	
Sev-2	Anti-depressants, 3-6 months via GP	
Sev-3	Anti-depressants, 3-6 months, with additional psychological support	
Sev-4	Combination therapy (medication and psychotherapy)	

See Lokkerbol et al. (2014a) for details regarding the underlying evidence base

Within the different diagnostic groups, this leads to 3, 3 and 6 different intervention pairs[1] that can be valued in terms of their potential to create a dominant health care configuration. In order to arrive at a ranking, each intervention pair is first categorized into 1 of the 6 scenarios as depicted in **figure 6.1**. Once it is known to which of the six categories each intervention pair belongs, we know which substitution leads to the highest net benefit. The health economic model is then used to simulate this substitution and calculate the additional net benefit, thereby visualizing uncertainty around the estimates using the cost-effectiveness plane and the cost-effectiveness acceptability curve. We use a WTP of €50,000 per DALY averted for illustrative purposes. A value of €20,000 could perhaps be deemed more appropriate, and could

---

[1] Mathematically, the number of pairs that can be chosen out of a group of 3 is 3! / (2!*(3-2)!), which equals 3. This indicates that from a group of 3, it is possible to select 3 different pairs. In the same way, the number of pairs that can be chosen out of a group of 4 is 4! / (2!*(4-2)!), which equals 6.

lead to less potential intervention pairs due to the constraint that the intervention that is invested in needs to have a CE-ratio below the WTP threshold value.

RESULTS

Each of the 12 intervention pairs was assessed. One of the intervention pairs was excluded because the CE-ratio of the intervention which should be scaled up was higher than the WTP threshold. The remaining 11 intervention pairs were valued and ranked in terms of their theoretical improvement potential by applying the algorithm as described earlier, see **table 6.2**.

Table 6.2: Improvement potential for major depression in the Netherlands

Intervention pairs	Coverage rate	Depression severity	Cost	Health	Net Benefit*	Additional Coverage
1 Anti-dep med + GP's assistant: Individual psychotherapy:	20% -> 38% 18% -> 14.75%	Severe	0%	+ 15%	+ 21%	14.75%
2 Anti-dep med + GP's assistant: Combination therapy:	20% -> 36% 16% -> 12.55%	Severe	0%	+ 12%	+ 16%	12.55%
3 Combination therapy: Individual psychotherapy:	16% -> 34% 18% -> 2.95%	Severe	0%	10%	+ 13%	2.95%
4 Anti-dep med + GP's assistant: Anti-dep med:	20% -> 38.50% 20% -> 0%	Severe	0%	+ 9%	+ 13%	-1.50%
5 Anti-dep med: Individual psychotherapy:	20% -> 38% 18% -> 15%	Severe	0%	+ 6%	+ 8%	15.00%
6 Anti-dep med: Combination therapy:	20% -> 36% 16% -> 12.82%	Severe	0%	+ 3%	+ 5%	12.82%
7 E-health supported: Individual CBT:	2% -> 18% 16% -> 12.33%	Moderate	0%	+ 2%	+ 3%	12.33%
8 E-health supported: Individual CBT:	2% -> 19% 17% -> 10.11%	Mild	-4%	+ 0%	+ 1%	10.11%
9 Individual CBT: Online psychotherapy:	16% -> 18% 2% -> 0%	Moderate	-2%	+ 0%	+ 1%	0.00%
10 E-health supported: Online psychotherapy:	2% -> 4% 2% -> 0.82%	Moderate	-2%	+ 0%	+ 1%	0.82%
11 E-health supported: Online psychotherapy:	2% -> 4% 2% -> 0.98%	Mild	-2%	+ 0%	+ 1%	0.98%

* defined as: (Net Benefit of the dominant health care configuration - Net Benefit of care as usual) / (Net Benefit of care as usual) * 100%

In the first intervention pair, containing interventions aimed at patients with severe depression, anti-depressant medication plus GP's assistant (intervention 1) is invested in at the expense of individual psychotherapy (intervention 2). We put a cap on the increase in coverage rate of intervention 1, equal to the coverage rate of intervention 2. This means that the coverage rate of intervention 1 can potentially

increase from its current rate of 20% to a maximum of 38%. The coverage rate of intervention 2 could potentially decrease to 0%.

In order to find the optimal coverage rates under the restriction of dominance, we first classify this intervention pair into one of the six categories as depicted in **figure 6.1**. Since i) the cost-effectiveness ratio of intervention 2 (mean: €12,681) is lower than the WTP of €50,000, ii) scaling up intervention 1 requires less budget than offering intervention 2 (C_1=€42 mln<C_2=€239 mln), and iii) health gains of scaling up intervention 1 are less than the health effects of offering intervention 2 (E_1=15,500 DALY<E_2=18,900 DALY), our situation corresponds to **figure 6.1f**. This means that the optimal way to adjust both interventions is to scale up intervention 1 to the full 38%, and to scale down intervention 2 up to the point where the overall change is budgetary neutral, which corresponds to lowering the coverage rate from 18% to 14.75%. In a final step, the health economic model can be used to generate the cost-effectiveness plane and the cost-effectiveness acceptability curve, to explore the uncertainty that comes with such a substitution.

Changing the intervention pair like this leads to a health care system that is expected to have a 15% increase in health effects under equal budget. However, as the coverage rate of intervention 2 decreases less than the increase in coverage rate of intervention 1, dominance in terms of cost-effectiveness requires a 14.75% increase in the total coverage in patients with severe depression. In any case, second-stage normative considerations may place restrictions on the extent to which these changes should be implemented. Therefore, the improvement potential as presented in **table 6.2** is only tentative and should be interpreted with caution. The results primarily serve the purpose of illustrating the algorithm.

CONCLUSION & DISCUSSION

Main findings

The main finding of this study is that virtually every intervention pair in a given health care system has the potential to improve the health care system overall in terms of costs and/or effects by investing and allowing for simultaneous disinvestment. Even having only relatively few interventions within a homogeneous target group, which is the case in our example, will lead to multiple improvement options. In reality, health care systems will consist of many more interventions, causing the number of intervention pairs and thus the available options to create a dominant health care configuration to increase substantially. Using existing interventions rather than new interventions to create a dominant health care configuration in terms of cost-

effectiveness has the advantage that implementation may be organisationally less challenging than offering an altogether new intervention, although disinvestment could potentially be scientifically, politically and ethically challenging (Pearson and Littlejohns 2007).

The algorithm results in strategic propositions for improving the cost-effectiveness of a health care system for consideration by policymakers and other stakeholders during a second-stage normative filtering process in terms of equity, feasibility, interchangeability, acceptability, appropriateness and strength of evidence. In the case of disinvestment, one could expect that it is difficult to change clinical practice when it comes to long-used, deeply entrenched technologies. Polisena et al. (2013) state that the introduction of a new technology, resulting in multiple available technologies for the same target group, could be an opportunity to initiate discussions regarding disinvestment.

Strengths and limitations

One of the strengths of this paper is the use of an algorithm that values intervention pairs within a health care system that can constitute a dominant health care configuration in terms of cost-effectiveness. This health economic substitution process requires a second-stage normative filtering process, where considerations other than cost-effectiveness should be taken into account.

An additional strength involves applying the methodology on the health care system for major depression in the Netherlands. Epidemiology is based on Dutch cohort data, clinical evidence is based on meta-analyses and RCTs, data on coverage rates are based on a health care provider focus group, data on compliance are based on a patient focus group, and full economic per patient costs of offering interventions is based on combining resource use with standard Dutch unit cost prices. In addition, the model makes all necessary information available in a dynamic form, allowing 'what – if' analyses. This could be of assistance to policy formulation and setting research agendas for innovation especially when the subject matter is intrinsically complex.

A strength of our paper involves taking the current intervention mix explicitly as a starting point when exploring routes to health care innovation using disinvestment. The WHO guideline on generalized cost-effectiveness analysis (Murray et al 2000) offers another route to innovation implicitly using disinvestment. In their approach, Murray et al (2000) do not take the current intervention mix as a starting point, but the "counterfactual of the null set of the […] interventions, i.e. the natural history of disease" (Murray et al 2000). However, ignoring the current

intervention mix and starting from 'a blank sheet', would only be relevant for countries with hardly any formal mental health care system in place, and where a new health care system could be designed from scratch. In our paper, using the current intervention mix helps in identifying improvement options that are more feasible from an implementation perspective, as changes to the health care system are likely to be less profound.

Another strength involves the transferability of our substitution algorithm, which transcends diagnostic context and could be applied to any health care system. In this context it is worth noting that we could apply the same algorithm on a different health economic model for different disorders, as described in Lokkerbol et al. (2014b). A limitation regarding transferability is that when investing or disinvesting has a non-linear impact on costs and effects (for example when looking at multi-year prevention models), our algorithm becomes computationally more challenging.

The results presented by the algorithm depict the improvement potential of health care substitution. Implemented fully, some of these results entail substantial changes regarding the health care system, as they involve large changes in coverage rates, or fully downscaling interventions. The results, however, indicate improvement directions. In **table 6.2**, it can be seen that when using intervention pair 9, the health care system could be improved by fully downscaling online psychotherapy (coverage rate from 2% to 0%), and correspondingly up scaling individual CBT (coverage rate from 16% to 18%). When a coverage rate of 0% for online psychotherapy is deemed inappropriate, then a change on a smaller scale could be implemented, for example by up- and downscaling the interventions with 0.2% instead of 2%. So even when the intervention pairs are not interchangeable for the entire target group, our algorithm still points out the direction in which smaller scale changes can be beneficial to the overall costs and/or effects of the health care system.

As with all modelling, the quality of our findings depends on the quality of the input data. However, data requirements for our method are particularly demanding, as high quality input is required from fields as diverse as epidemiology, economics and clinical evidence from randomized trials. A lack of high quality evidence leads to uncertainty around outcome parameters, which, although explicitly taken into account by the sensitivity analysis of the health economic model, is likely to impact the results cf. Lokkerbol et al. (2014a).

Our results are limited to presenting a list of intervention pairs that have the potential to constitute a dominant health care configuration. Adjusting the health care system to actually achieve such a dominant configuration is a challenge in its own right, such as selecting that intervention pair that will be accepted by health care

users and is regarded appropriate by health service providers; or bringing the health care system from one equilibrium to another, possibly requiring substantial investments or overcoming political barriers that are likely to occur in the context of disinvestment. In that sense, our improvement options do not only entail a health system change, but are likely to require clinical behaviour change as well, as it requires health care providers to offer a different mix of existing interventions.

Disinvestment is an essential component of the algorithm presented, yet the process of disinvestment is associated with difficult challenges (Pearson and Littlejohns 2007). Tsourapas and Frew (2011) show in their literature review on priority-setting using Programme Budgeting and Marginal Analysis (PBMA), that this can lead to disinvestment and redirected resource allocation. At the same time, Polisena et al. (2013) emphasize in their systematic review on case studies on disinvestment and resource allocation, that in most case studies, it is unknown whether recommendations regarding disinvestment were implemented, and if so, what the impact was on patient care, health services delivery and cost to the health care system. Additional research is needed to investigate to what extent it is feasible to substitute one intervention for another. Also, more information is needed to gain insight in the possible adverse effects of disinvestment and how these could be minimised.

Our substitution algorithm should be used for homogeneous target groups. The more heterogeneous the patients within a diagnostic target group, the more likely it is that intervention pairs with improvement potential do not pass the second-stage filtering process as substitution will not be considered feasible. Substitution is further complicated as some interventions are known to be cost-effective only for subgroups of the population (Chandra and Skinner 2012), meaning that even within homogeneous target groups, strength of evidence regarding the true interchangeability of interventions should be explicitly investigated.

Another limitation is that our algorithm uses point estimates of costs and effects as a starting point for valuing and ranking the different substitutions. Uncertainty is not taken into account until after the ranking, when the health economic model is used to simulate the substitutions, explicitly taking uncertainty into account. For policymakers with a pronounced preference for low uncertainty this could alter the ranking.

A final limitation is that our substitution algorithm only considers changes to the health care system resulting in a health care system that ends up in the lower right quadrant of the ICER-plane (based on point estimates). Alleviating the restriction that the new health care system needs to have at least as much health effects and cost no more than care as usual, could lead to more improvement

potential. However, we imposed this restriction as it is deemed to be more consistent with the current economic and political tide.

Implications

In times of the economic downturn and entailing budget cuts in health care, it is important for policymakers in the health care sector to design strategies that are likely to have a beneficial impact on the cost-effectiveness of the health care system as a whole. Our health economic model and algorithm point out which existing intervention pairs have the potential to constitute a dominant health care configuration via health care substitution. Although many of the intervention pairs will be deemed inappropriate during a second-stage filtering process, and although disinvesting interventions could involve overcoming political barriers, we see great value in offering policymakers an additional set of propositions for innovating the health care system to increase health gains in the population for a similar or smaller health care budget.

Summary & General Discussion

Rationalization of Innovation: summary and general discussion

The previous chapters in this thesis covered 1) the need for a cost-effective mental health care system, 2) established methods for assessing cost-effectiveness, and 3) new approaches to improving cost-effectiveness. This summary returns to these topics.

The need for a cost-effective mental health care system

The need for a cost-effective mental health care system is driven by a large disease burden on the one hand and scarce resources on the other. Mental disorders have a prominent position on the global ranking of (non-fatal) disease burden (Vos et al 2012). In addition, health care only manages to partially reduce the disease burden due to mental disorders at population level (Andrews et al 2004; Chisholm et al 2004a). At the same time, available resources in the form of health care budgets and human resources are under pressure. Health care expenditure as a percentage of GDP has been steadily increasing over the last decade (OECD Stats) and demographic developments such as aging populations are expected to lower the number of people in the working population relative to the number of older people (United Nations 2013), who are generally associated with on average higher health care demand (Wong et al 2012). These developments emphasize the need for more cost-effective ways to alleviate disease burden under resource constraints, even more so when taking developments like the economic downturn into account.

Improving cost-effectiveness generally starts with acknowledging the disease burden and the economic consequences of disorders. Both individual as well as population disease burden are important for deciding where health care budgets should be directed. From a health economic perspective, it is important to understand that highly prevalent disorders such as social phobia could put just as much pressure on population health as individually highly disabling but less prevalent disorders such as schizophrenia, as a modest disease burden on individual level combined with high prevalence and a long duration could amount to a larger population disease burden than a more severe individual disease burden applying to a smaller group of people would.

Chapter 1

A typical starting point for health care innovation is therefore to describe both the individual and population level disease burden. Chapter 1 describes the non-fatal disease burden due to mental disorders in the Netherlands, at both the individual and population level. It is shown that from a population level, it is not the individually disabling disorders such as bipolar disorder or schizophrenia that drive disease burden, but simple phobia, social phobia and dysthymia which are all characterised by individual disease burden combined with a high prevalence and long disease duration.

Established methods for assessing cost-effectiveness

There are various methods for economic evaluation that are commonly used to evaluate the (relative) cost-effectiveness of interventions with respect to a comparator condition. Broadly these methods can be categorized in trial-based economic evaluation and economic evaluation using decision analytic modelling, see Drummond et al. (2005) and Briggs et al. (2006) for more detail. These methods are commonly used methods today, and their application is presented in chapters 2 – 4.

Chapter 2

Chapter 2 presents a cost-effectiveness analysis of a trial with four arms concerning the treatment of patients with (symptoms of) depression and/or anxiety. The trial investigates the effectiveness of an online intervention with varying degrees of therapeutic support in reducing symptoms of anxiety and depression. Data on costs and effectiveness were gathered over a period of 12-months. By performing a multi-arm bootstrapping method, we found that the intervention condition with no support has a high probability of having a more favourable cost-utility ratio than the active control condition of non-specific chat or email support. Although high dropout rates make our conclusions only tentative, it is hopeful that higher levels of support do not necessarily seem to lead to the most favourable outcomes, as availability of human resources in health care can be expected to become an increasingly important constraint in the future.

Chapter 3

Chapter 3 presents a decision analytic model that assesses the population-level effect of adding online interventions onto a health care system for alcohol use disorders. For this purpose, the health care system for alcohol disorders in the Netherlands was modelled for the target groups defined in terms of abstinence, moderate drinking, heavy drinking, hazardous use, harmful drinking and alcohol dependence, following

the terminology introduced by the World Health Organization (1994). The health economic modelling shows that care as usual is associated with a benefit-to-cost ratio of €1.08, which can be improved by adding online interventions yielding a health care system with an overall benefit-to-cost ratio of €1.62. Results relate to the short-term of one year and assume a steady state after full implementation of the alternative health care system (i.e. with e-health interventions added). This explicitly disregards implementation costs (i.e. training health care professionals) as well as the time required to move the health care system from the old to the new equilibrium.

Chapter 4

Improvements in terms of cost-effectiveness often require additional health care budget, as new interventions are added to the current intervention mix. Chapter 4 investigates how health care substitution can compensate for steadily increasing health care expenditure due to adding new interventions to the current package of interventions. In this chapter we look at the amount in which coverage rates of existing interventions need to decrease in order to improve health effects without increasing the budgetary ceiling.

To that end, we used Markov modelling to assess the impact on cost-effectiveness when adding preventive telemedicine to the health care system for major depression in the Netherlands. The Markov model investigates the longer-term (five-year) impact in terms of costs and health effects. To this end, the epidemiology of major depression in the Netherlands was modelled, taking into account yearly incidence and prevalence rates while distinguishing between subclinical, mild, moderate, severe and chronic depression. The current mix of interventions, where all interventions are aimed at treatment, is associated with a benefit-to-cost ratio of €1.30, meaning that every euro invested is expected to generate €1.30 in terms of health benefits. Interventions aimed at prevention are associated with a more profitable benefit-to-cost ratio of €1.60 and are therefore considered more cost-effective than the current intervention mix. When adding realistic levels of preventive interventions to the current intervention mix with, the overall benefit-to-cost ratio increases from €1.30 to €1.32 over a five-year period. The results do not take implementation costs into account.

Adding prevention increases the overall required health care budget with 7%. This increase is relatively low as successful prevention saves treatment costs at a later stage. When downscaling the interventions aimed at treatment in order to arrive at a budget-neutral scenario, overall budget does not change, while offering relatively more cost-effective interventions (prevention) results in more health effects.

This again increases the benefit-to-cost ratio from €1.30 to €1.32, but this time with no change in overall budget. As budgets are not easily increased in times of economic downturn, it is important to explore different approaches to improve the cost-effectiveness of mental health care.

After five years of prevention, a new equilibrium of patients needing treatment is not yet achieved. Prevention decreases the amount in which treatment is needed, so as long as a new epidemiological equilibrium is not achieved, each additional year of prevention will be associated with an overall more favourable benefit-to-cost ratio.

It is worth mentioning that for this thesis the Markov model presented in chapter 4 was updated with respect to the model presented in the publication of Lokkerbol et al. (2014a) in terms of the underlying epidemiology, which is now based on Nemesis-2 which employed the CIDI/DSM-IV (de Graaf et al 2012b) and no longer on Nemesis-1 which was based on CIDI/DSM-III-R. The new version of DepMod is also updated with respect to the evidence-base regarding prevention (van Zoonen et al 2014). The conclusions, however, are similar: prevention is cost-effective.

New approaches to improving cost-effectiveness
Health economic modelling is an evolving field. As our understanding of health economic modelling increases, it becomes possible to pursue different routes to innovation. Chapters 5 and 6 present two new approaches to improving the cost-effectiveness of the health care system, that is, approaches that are different from adding new, presumably more cost-effective interventions to the health care system, but for example incorporate mixtures of investing and disinvesting in competing interventions for the same diagnostic target group. Furthermore, these approaches differentiate themselves from commonly used approaches in the sense that they do not investigate one specific, new intervention. Rather, they investigate the range of existing health care interventions in order to provide a list of options to improve the health care system overall. These approaches are therefore best seen as algorithms that systematically scan the current health care system in order to identify options to improve cost-effectiveness. The list of options is meant to provide policy-makers with input regarding system innovation, which can be assessed from normative perspectives other than cost-effectiveness, such as equity, acceptability, appropriateness, feasibility and strength of evidence.

Chapter 5

Chapter 5 investigates the potential for improvement of mental health services in the primary care setting in the Netherlands by systematically scanning the impact on the health care system of local improvements in either coverage, adherence, effectiveness or costs of each target group – intervention combination. Each of these parameters plays a role in the cost-effectiveness of the overall health care system. Scanning the target group – intervention – parameter combinations for their ability to leverage the overall cost-effectiveness of the health care system enables innovation to be geared towards those areas where it is expected to have the largest impact. As the potential of an intervention to improve the cost-effectiveness of the health care system is largely determined by the absolute size of the target group it serves, the list of interventions with the most improvement potential when improving coverage, adherence, effectiveness or cost is mostly dominated by highly prevalent interventions, such as pharmaceutical interventions.

Increasing the coverage rates of interventions in order to improve the cost-effectiveness of the health care system can be considered reaping the benefits of already cost-effective interventions. Increasing the adherence or the effectiveness of interventions, or decreasing the intervention's costs can be considered favourable investment strategies meant to unlock the potential for improvement in these areas. As the proposed improvements are mostly conjectural, critically examining these improvements from a broader perspective is again necessary.

Chapter 6

Chapter 6 takes a health care substitution approach to improving the cost-effectiveness of a health care system by analysing the impact of investing in relatively cost-effective interventions while simultaneously disinvesting in relatively cost-ineffective interventions. An algorithm is developed that systematically scans the potential of each intervention pair to constitute a health care system that generates at least as much health effects under equal or less budget (a dominant health care configuration), by investing in one intervention and simultaneously disinvesting in the other. It is shown that in general there are as many possibilities to create a dominant health care configuration, as there are different intervention pairs in a health care system. As the number of possible intervention pairs increases substantially when the number of interventions increases, this algorithm can provide policymakers with a long list of options for improvement that do not require investment in new interventions.

The algorithm is applied to major depression in the Netherlands, using the same health economic (Markov) model as in chapter 4. It is shown that there are 11 different intervention pairs with the potential to create a more efficient health care system. The algorithm is not restricted to specific (e.g. Markov) models or the mental health care setting per se, but is a generally applicable tool. As outcomes do not involve an increase in budget but do involve disinvesting, this approach could shift the political discussions from financial to feasibility and ethics.

Broader perspectives

For the economic evaluations presented in this thesis, and for economic evaluations in general, it often holds that proposed innovations are expected to lead to a more cost-effective health care system. Of course, these innovations should not be implemented straight away, as health economic modelling can never be seen as an autopilot for innovation. The economic evaluations in chapters 2 – 6 approach the health care system from a cost-effectiveness point of view. Whether proposed changes are desirable from a societal perspective depends on other factors as well. Therefore, a second-stage filtering process should follow the results of health economic modelling, where preferably criteria such as equity, appropriateness, acceptability, feasibility and strength of evidence should be taken into account. This normative second-stage filtering process is desirable for most health economic evaluations, but is even more needed for the algorithms discussed in chapters 5 and 6, where many possible improvements are listed requiring second-stage filtering. In order to increase alignment with criteria other than cost-effectiveness, the process of health economic modelling is preferably guided by input from the clinical, patient and policy perspective. The development of multidisciplinary clinical guidelines, where clinicians and patients can actively participate in the development of the (conceptual) framework of the health economic model, is a good example of this.

General Discussion

This thesis presents five papers using different approaches to evaluating and improving the cost-effectiveness of mental health care in the Netherlands, preceded by a paper presenting the overall (non-fatal) disease burden due to mental ill-health in the Netherlands. This final chapter reflects on the findings, discusses the implications and suggests directions for future research.

The underlying theme of this thesis is the application of economic evaluation in mental health care. Health economic modelling involves combining health sciences with economics in an attempt to take both perspectives into account for decision and policy-making. Health and economics often require a deliberate effort to be combined into a single policy-making perspective (as pursued by the extra-welfarism school of economic thought), as both fields tend to value different outcomes. Solely pursuing health interests can conflict with economic interests and vice versa. Both the health and the economic perspective have their value. However, not evaluating both perspectives simultaneously is likely to result in one perspective dominating the decision-making process, which in turn could lead to sub-optimal decision-making.

Loosely speaking, providing qualitatively good health care (aimed at promoting, protecting and restoring population health) is the ultimate goal of health care innovation. Whereas economics is concerned with the conditions (resources) needed to provide this health care, health-economics pursues providing the best possible health care (i.e. sustaining population health in an optimal way) given the available resources. Sub-optimal use of resources (which can be interpreted as wasteful use of valuable human and financial resources) does not contribute to sustainable population health.

In times of economic prosperity, economic constraints are less restricting and it is possible for health care and health care research to prosper. However, during the current economic downturn, the economics of health care has become more dominant, with an increased interest in for example budget impacts of health care innovations. This process has been accompanied by budget cuts and major health care reforms, which could for example lead to decreasing coverage of health care services or increasing per patient user charge (Karanikolos et al 2013).

Towards resource-efficient health care
The relatively dominant position of economics in health care is not a bad thing in itself. Potentially it is a good thing, as taking into account financial realities is a prerequisite for sustainable health care both for the current and future generations. However, there are multiple ways in which economics can join the debate. To simplify, the economics perspective could introduce a focus on budget cuts without taking health effects into account, which is likely to save budget in the short term but can have a negative impact on population health (Karanikolos et al 2013). Preferably, however, the economics perspective could introduce a focus on resource-efficient health care, where all efforts in health care are judged by their ability to contribute

to health and their required human resources and budgets. This requires decisions regarding health care to be made using an explicit decision-making framework.

It is likely that the economic situation will improve at some point, as economic prosperity tends to come in waves. However, with an eye to the more structural demographic transition, future economic prosperous times should not be seen as signs that health care reforms are not necessary, but as opportunities to set up more resource-efficient health care systems, not only taking into account the financial budgets but also the (limited) human resources that will restrict the width and breadth of our future health care. It is inevitable that ethical discussions regarding the delivery of health care will become increasingly important in the coming decade. It is only by the extent in which we manage to create a resource-efficient health care system, that we can keep those ethical dilemmas at bay.

Limitations

There are several limitations to this thesis. Combining evidence from different fields into health economic simulation models allows for an insightful macro-perspective on health, but comes at the cost of substantial data requirements and the need to make assumptions. Uncertainty arises when input parameters cannot be based on firm evidence, due to unknown longer-term impacts of changes in the health care system, and due to epidemiological uncertainties regarding the pathogenesis and (natural) course of disorders. Moreover, outcomes should always be interpreted in terms of the quality of the inputs used, the time-horizon in which results are presented, and the perspectives (society, health care sector, etc.) on which the results are based. It is important to realize that health economic modelling studies do not provide instant solutions, but provide guidance in the wider context of the decision-making process (Niessen et al 2012).

Another limitation to the health economic modelling studies included in this thesis, is that the perspective is by and large restricted to costs and clinical effects. This means that results of health economic modelling studies should ideally be followed by a normative second-stage filtering process, where perspectives other than cost-effectiveness, such as equity, medical ethics, feasibility, acceptability, appropriateness and strength of evidence are taken into account, in order to have value for policy-making (Vos et al 2010). Outcomes of health economic modelling are thus best seen as well-informed, rational options for improvement, thereby providing a starting point for policymakers who then need to take into account second-stage filtering criteria (Berghmans et al 2004; Mihalopoulos et al 2011b).

Next to being restricted by and large to economic costs and clinical effects, most studies in this thesis are restricted to cost-effectiveness within the health care sector. It would be interesting to get a better idea on how investing in mental health could impact on other sectors, for example education, or productivity in the long-term. In addition, learning how the return-on-investment within health care relates to the return-on-investment in other sectors could contribute to finding an optimal budget allocation across sectors.

Another limitation is that health economic models in general present the improvement potential in terms of cost-effectiveness (and/or budget impacts) of new interventions specifically, or alternative health care scenarios in general, which is commonly evaluated comparing the current health care system with an alternative health care system, when each health care system is assumed to be in a 'steady state' (health economic equilibrium). By assuming each health care system to be in its steady state, we explicitly disregard the time, money and effort required to move the health care system from the current equilibrium to a new one. The investment required to accomplish such a transition could be substantial, for example when the new health care system requires capacity building in the form of training of health care professionals on a large scale, or when the change in the health care system is substantial and can be expected to take years before fully implemented and adopted by all stakeholders. Our health economic models do not take into account these transition costs, but merely focus on the comparative cost-effectiveness of the new and old health care systems when both systems are in a steady state (similar to the approach described in for example Vos et al. (2010)).

A final limitation is that this thesis is work in progress. Methods to evaluate health care systems develop along with developments in society. In times of economic prosperity, health economic evaluation will put less emphasis on budgetary constraints than in times of economic downturn. Also, as the number of available interventions keeps growing steadily, health care financiers need to be more critical towards the interventions they reimburse. On the one hand this results in an increasing interest in acceptable, affordable and effective care by health care financiers. On the other hand, this creates the incentive for health care providers to understand the way in which health care financiers evaluate interventions. This leads to an increasing interest among health care providers to speak the same language as the more economically oriented health financiers with whom they have to negotiate tariffs and reimbursements.

Future directions

There are several issues that should be incorporated in the field of health economic modelling in the near future. First, apart from monetary resources, human resources should become a fully integrated part of health economic evaluations, by letting required human resources be an equally important and constrained input (next to costs) in health economic evaluations. There is little scientific challenge in incorporating such an input (commodity), as resource use in health care is commonly measured in most cost-effectiveness analyses (see for example Smit et al. (2006b)). Incorporating human resources as an input could add an additional dimension to for example the algorithm presented in chapter 6. For each intervention pair considered, one could look at the human resources needed when investing in an intervention, as well as the human resources that become available when disinvesting in an intervention. Ideally, this would help to improve the allocative efficiency of investing and disinvesting in intervention pairs such that the health care system overall becomes more effective, less costly and less demanding with respect to human resources. Incorporating this approach gives the opportunity to consider the resource that is most constraining, be it financial or human.

An additional challenge with regard to human resources as opposed to financial resources is that the former is not as freely transferable as the latter. Money, when not spent on health care, could be used elsewhere. People trained for a health care profession are not as easily employed elsewhere or trained to offer a different type of intervention. Changes in the health care system aimed to improve the resource-effectiveness could therefore be accompanied by short-term market inefficiencies, entailing a real cost on the system level. Future research could help in our understanding of these effects when changing the health care system.

More research is needed regarding the dynamics of the transition from one (health care) equilibrium to the other. Discussions around health care system improvement would benefit from increased understanding on several aspects, such as whether the term 'equilibrium' is justified in the first place, whether changes in costs and effects are proportional to the change in the underlying intervention mix or whether this relation is less straightforward (e.g. non-linear), and how the need for re-schooling and re-employing affects the transition and implementation costs.

Thirdly, disinvestment from ineffective or inappropriately applied practise is growing as a priority (Elshaug et al 2008), yet there are challenges that limit the use of such a strategy. As interesting directions for health care improvement could be based on well-balanced combinations of investment and disinvestment, future research should increase our understanding of the (adverse) effects of disinvestment

and ways to compensate for this, such that a new potential area for improving the health care system can be unlocked.

Furthermore, health economic modelling would benefit from a more profound understanding of the inter-sectorial costs and benefits of health care interventions. In the Netherlands, the full economic cost prices within the sectors of education and the criminal justice system were recently mapped (Drost et al 2014), providing the opportunity to explicitly incorporate inter-sectorial costs and benefits in a methodologically consistent approach. By modelling the wider impact on society, the decision-making process will be better informed, which could potentially open up the road to re-distributing costs and benefits of interventions across sectors to try to create financial constructions where every stakeholder can benefit.

Next, in order to create real policy impact, health economic evaluations should be further integrated in the decision making process of policymakers. Health economic evaluation can provide valuable, complementary information that could inform the decision-making process. However, carrying out an economic evaluation too much in a stand-alone fashion, removed from the specific interests of decision makers, could result in missing valuable aspects important to policymakers, failing to reach the relevant decision makers, or otherwise limiting the use of the economic evaluation in decision-making (Hoomans et al 2007; Eddama and Coast 2008; Niessen et al 2012). The development of multidisciplinary clinical guidelines (for example NICE Guidelines 2011; 2014) is one particular setting in which model development can benefit from diverse perspectives. These settings are helpful in themselves, as they could have a positive impact on patient adherence to interventions and clinicians' adherence to guidelines, which in turn can be expected to impact positively on cost-effectiveness. In general, health economic evaluations could increase their added value by actively involving stakeholders and incorporating their interests (e.g. by identifying opportunities for multi-stakeholder win-win solutions).

Also, to create a greater impact on clinical practice, the possibility of integrating health economic evaluation within clinical practice should be investigated. This would require a mechanism to translate insights based on a mostly internationally oriented (macro-level) evidence base, into valuable input for the patient-therapist (micro-level) setting. As this micro-level inherently differs from the macro-level, such a mechanism should facilitate continuous learning at the micro-level, and should therefore involve some form of outcome monitoring in order to capture the effect of implementing evidence-based health care in the local, patient-therapist, setting. However, evaluating patient-level process data requires potentially new types of statistical methods, not commonly used in the field of mental health care. Valuable lessons could be learned from sectors where it is more common to

create an environment of continuous learning based on process data, such as industry, financial services or somatic health care (Lokkerbol et al 2011; 2012a; 2012b; de Mast et al 2012), whereas micro-simulation (i.e. discrete event simulation) might be the way forward in health-economic modelling of micro-economic processes and outcomes.

Needless to say, a patient in mental health differs from a client at a bank, or a product in industry, or even a patient in somatic care. Yet, with some adjustments, statistical techniques developed for monitoring processes, such as statistical process control (Shewhart 1939) are likely to have value in the mental health sector as well. This approach could introduce a shared (patient – therapist) perception and provide feedback on the therapeutic process. As this would possibly require a different role for both the therapist and the patient, it is important to monitor whether this introduces undesirable (adverse) side effects. Future research should explore how economic evaluations can contribute more at the patient-therapist level. Because, even though health economic modelling, through synthesizing evidence in the fields of epidemiology, clinical effectiveness and economics, is likely to start in the domain of science and may then contribute at the macro-level, the ultimate challenge lies in contributing at the micro-level where the actual health gains and costs are generated.

Conclusion

In this thesis we showed different methods for evaluating the cost-effectiveness of health care systems, thereby providing input to make health care more cost-effective. Much is already known about the methodologies underlying health economic evaluation (for example Murray et al 2000; Drummond et al 2005; Mauskopf et al 2007; Siebert et al 2012; Husereau et al 2013). The next challenge therefore seems to be to align health economic evaluation with the current and often local economic, epidemiological and demographic realities. Efforts aimed at improving cost-effectiveness by adding additional, presumably cost-effective interventions to the current intervention mix, though informative, are likely to clash with budgetary ceilings in the context of the current economic downturn, i.e. times when it is difficult to increase budgets. On the other hand, health economic evaluations investigating the optimal (re-)configuration of the current intervention mix (see for example Murray et al. (2000)) could potentially entail large scale health care reforms, which is likely to be difficult politically.

An intermediate way is to take the current intervention mix as a starting point in order to increase realism and to better relate to policymakers' realities, and then propose improvements that do not exceed the current budget and are therefore

more in line with the current economic realities. Chapter 6 shows that every health care system is likely to have possibilities for improvement satisfying these conditions.

Optimizing the health care system under fixed budgets using combinations of investment and disinvestment opens up a new array of possibilities for improvement that can help to set up a sustainable health care system. However, it is likely that political opposition will occur when existing interventions are being targeted for downscaling. Finding balanced and acceptable ways to downscale interventions will be a new challenge altogether. However, in certain countries, such as the Netherlands, there are stakeholders like health care insurance companies that negotiate budgets with health care providers. Decreasing budgets can occur as a result of these negotiations, thus providing a setting for health care substitution. It is of course crucial for health effects to be specifically taken into account in such decisions. As disinvestment is associated with many and diverse challenges, future research, involving all relevant stakeholders, should aim to minimize the undesirable effects of disinvestment, such that despite the economic downturn and demographically driven constraints in the labour market for health care, a health care system can deliver its promise of maintaining population health in a sustainable way.

To summarize, this thesis investigated the need for a cost-effective mental health care system, employed established methods for assessing the cost-effectiveness of interventions, and developed new approaches to improving the cost-effectiveness of health care systems.

The need for a cost-effective mental health care system is considered to be high due to the relatively prominent position of mental disorders on the burden of disease rankings, the gap between the disease burden in the population and the disease burden averted by treatment, and the demographic developments that can be expected to further complicate the sustainability of health care. Established methods of economic evaluation contribute to the knowledge base needed to arrive at a more cost-effective mental health care system, but often do not take budget constraints into account. It requires different methods of economic evaluation to explore options to improving the cost-effectiveness of health care under explicit budget constraints. This thesis shows that it is possible to develop such methods. However, as improving the cost-effectiveness under explicit budget constraints is likely to involve disinvestment, this requires a paradigm shift both in policymaking –with the relevant stakeholders– and in the research field.

Implications

Policymakers need to start considering the possibility of using disinvestment to fund new, more cost-effective interventions, and the consequences this has for the decision-making process. Making difficult decisions, next to involving relevant stakeholders, requires a 'fair decision-making process'. Policymakers may need to consider what constitutes a 'fair decision-making process' in the eyes of the relevant stakeholders. A framework such as Accountability for Reasonableness (Daniels and Sabin 1997, 1998; Martin et al 2002; Friedman 2008), defining 'fairness' in terms of the publicity condition (rationales for priority setting must be publicly available), the relevance condition (these rationales must be considered by 'fair-minded' people relevant to priority setting in the considered context), the appeals condition (there must be a means to appeal decisions or rationales) and the enforcement condition (there must be a means to ensure the previous conditions) (Martin et al 2002), could be helpful in this context.

Successful investment / disinvestment decisions resulting from fair processes can only be possible when the required evidence is available. The research field has to provide the evidence needed to enable decision-makers to make balanced decisions. In the context of disinvestment, this requires a subtle change in research focus, from analyzing the (cost-)effectiveness of interventions within diagnostic categories to considering the (cost-)*in*effectiveness of interventions within *sub*groups of diagnostic categories. An obstacle to successfully implementing disinvestment is that there are rarely data from rigorous studies to demonstrate ineffectiveness across all types of patients (Pearson and Littlejohns 2007). By getting a better understanding regarding the subgroups of patients where interventions are indeed ineffective, well-informed disinvestment in order to fund more cost-effective interventions can be stimulated.

As disinvestment is considered to be a highly local process (Pearson and Littlejohns 2007), health care providers are essential in bridging the gap between research and practice. Next to representing one of the disciplines that should be engaged for a priority setting process to be considered 'fair' (see for example Martin et al 2002), health care providers are the stakeholders best positioned to turn the improvement of cost-effectiveness from one-time projects (for example through the development of multidisciplinary guidelines), into a continuous process (for example through monitoring treatment effects, thereby continuously adding to our understanding of how improvement suggestions unfold after their implementation, in terms of the effectiveness of an intervention within subgroups of patients, and of the knowledge gaps that should be considered in future research).

Health care innovation involving disinvestment is certainly not an easy solution, as this is likely to tap right into difficult ethical issues. Therefore, such an approach can only succeed as a joint effort by policymakers, health care users, health care providers and researchers. Even though disinvestment is not easy and is likely to involve both winners and losers: in the end we are best served by evidence-based, acceptable, appropriate and sustainable health care. Or, as Donaldson et al. (2000) put it: "To minimize harm, 'rational disinvestment' is the only logical way forward".

Samenvatting

Rationalisatie van innovatie:
de rol van gezondheidseconomische
evaluatie in het verbeteren van
de efficiëntie in de geestelijke
gezondheidszorg

De noodzaak van een kosteneffectieve mentale gezondheidszorg

De noodzaak voor een kosteneffectieve mentale gezondheidszorg wordt gedreven door een grote ziektelast enerzijds en beperkte middelen anderzijds. Mentale stoornissen hebben een prominente positie op de ranglijst van wereldwijde veroorzakers van (niet-fatale) ziektelast (Vos ea, 2012). De gezondheidszorg slaagt er echter slechts ten dele in om de ziektelast van mentale stoornissen op populatieniveau te verminderen (Andrews ea, 2004; Chisholm ea, 2004a). Tegelijkertijd staan de beschikbare middelen in de vorm van budgets en zorgprofessionals onder druk. Zorguitgaven, uitgedrukt als percentage van het BBP, zijn de afgelopen tien jaar gestegen (OECD Stats) en demografische ontwikkelingen als de vergrijzing leiden naar verwachting tot een relatieve daling van het aantal mensen in de beroepsbevolking ten opzichte van het aantal ouderen (United Nations, 2013), terwijl ouderen gemiddeld genomen juist meer zorg nodig hebben (Wong ea, 2012). Deze ontwikkelingen benadrukken de noodzaak om de kosteneffectiviteit van de mentale gezondheidszorg te verbeteren, op een manier waarbij expliciet rekening wordt gehouden met de beschikbare middelen. Deze noodzaak wordt nog eens versterkt door ontwikkelingen zoals de economische crisis.

Het nastreven van verbeterde kosteneffectiviteit begint over het algemeen met het in kaart brengen van de ziektelast en de economische gevolgen van stoornissen. Ziektelast op individueel en op populatieniveau zijn belangrijk in de besluitvorming omtrent de besteding van zorgbudgets. Vanuit een gezondheidseconomisch perspectief is het belangrijk om te realiseren dat veel voorkomende stoornissen zoals sociale fobie net zoveel druk op de gezondheid van een populatie kunnen uitoefenen als individueel sterk belastende, maar minder vaak voorkomende stoornissen zoals schizofrenie. Een bescheiden ziektelast op individueel niveau, wanneer deze wordt gecombineerd met een hoge prevalentie en een lange ziekteduur, kan op populatieniveau meer ziektelast veroorzaken dan een grote individuele ziektelast die betrekking heeft op een kleinere groep.

Hoofdstuk 1

In hoofdstuk 1 werd de niet-fatale ziektelast als gevolg van mentale stoornissen in Nederland in kaart gebracht, zowel op individueel als op populatieniveau. Op populatieniveau wordt de meeste ziektelast niet veroorzaakt door de individueel sterk belastende stoornissen zoals bipolaire stoornis of schizofrenie, maar juist door specifieke fobie, sociale fobie en dysthymie, die allen worden gekenmerkt door een individuele ziektelast die wordt gecombineerd met een hoge prevalentie en een lange ziekteduur.

Gevestigde methoden om kosteneffectiviteit te evalueren

Er zijn verscheidene methoden voor economische evaluaties die vaak worden gebruikt om de (relatieve) kosteneffectiviteit van een interventie af te zetten tegen een vergelijkingsconditie. Grofweg kunnen deze methoden worden onderverdeeld in op trials gebaseerde economische evaluaties en economische evaluaties die gebruik maken van gezondheidseconomische modellen, zie Drummond ea (2005) en Briggs ea (2006) voor meer informatie. Dit zijn de meest gangbare methoden, welke worden gedemonstreerd in de hoofdstukken 2, 3 en 4.

Hoofdstuk 2

In hoofdstuk 2 werd een kosteneffectiviteitsanalyse beschreven, gebaseerd op een trial met vier condities gericht op patiënten met (symptomen van) depressie en/of angst. De trial onderzoekt de effectiviteit van een online interventie met verschillende niveaus van begeleiding in het verminderen van angst- en depressiesymptomen. Data omtrent kosten en effecten werden verzameld gedurende 12 maanden. Een multi-arm bootstrapping methode liet zien dat de interventie zonder begeleiding een hoge kans had op een betere kostenutiliteitsratio dan de controleconditie met niet-specifieke begeleiding via chat of per email. Door de hoge uitval kunnen slechts voorzichtige conclusies worden getrokken. Desalniettemin is het hoopvol dat meer intensieve vormen van begeleiding niet noodzakelijkerwijs lijken te leiden tot betere uitkomsten, gegeven de verwachte druk op de beschikbaarheid van zorgprofessionals in de toekomst.

Hoofdstuk 3

In hoofdstuk 3 werd een gezondheidseconomisch model gepresenteerd waarin de populatie-effecten worden geschat van het toevoegen van online interventies aan de bestaande zorg voor alcoholstoornissen. Hiervoor werd het zorgsysteem voor alcoholstoornissen in Nederland gemodelleerd voor de doelgroepen: 'abstinence', 'moderate drinking', 'heavy drinking', 'hazardous use', 'harmful drinking' en 'alcohol dependence', in navolging van de terminologie zoals deze door de WHO wordt gebruikt (World Health Organization, 1994). Het gezondheidseconomisch model laat zien dat de huidige zorg een opbrengst-kostenratio heeft van €1,08, en dat deze verbetert tot €1,62 wanneer online interventies worden toegevoegd aan het zorgsysteem. De resultaten hebben betrekking op de korte termijn van één jaar en veronderstellen dat het alternatieve zorgsysteem (met online interventies) volledig geïmplementeerd is en een stabiel nieuw evenwicht heeft bereikt. Hiermee worden implementatiekosten (bijvoorbeeld het trainen van zorgprofessionals) dus expliciet

buiten beschouwing gelaten, alsmede de tijd die nodig is om het zorgsysteem van het oude naar het nieuwe evenwicht te verschuiven.

Hoofdstuk 4

Verbetering in termen van kosteneffectiviteit vereisen vaak additioneel zorgbudget, wanneer deze verbetering het gevolg is van het toevoegen van nieuwe interventies aan de huidige interventiemix. In hoofdstuk 4 wordt onderzocht hoe zorgsubstitutie kan compenseren voor stijgende zorgbudgets als gevolg van het toevoegen van nieuwe interventies aan de bestaande zorg, door te onderzoeken in welke mate het aanbieden van interventies gericht op behandeling zou moeten afnemen om gezondheidswinst te bewerkstelligen zonder overschrijding van het budgettaire plafond.

Hiervoor werd in dit hoofdstuk met behulp van een Markov model onderzocht wat de impact is van het toevoegen van preventieve online interventies op de kosteneffectiviteit van het zorgsysteem voor depressie in Nederland. Het model onderzoekt de lange termijn impact (vijf jaar) in termen van kosten en gezondheidseffecten. Hiervoor werd de epidemiologie van depressie in Nederland gemodelleerd, gebaseerd op de jaarlijkse incidentie en prevalentie van depressie, waarbij onderscheid werd gemaakt tussen subklinische, milde, matige, ernstige en chronische depressie. De huidige interventiemix, die volledig gericht is op behandeling, is geassocieerd met een opbrengst-kostenratio van €1,30, wat betekent dat elke geïnvesteerde euro naar verwachting €1,30 aan gezondheidsgerelateerde waarde genereert. Interventies gericht op preventie zijn geassocieerd met een hogere opbrengst-kostenratio van €1,60 en kunnen daarom worden beschouwd als meer kosteneffectief dan de huidige interventiemix. Wanneer de huidige interventiemix wordt uitgebreid met een realistische hoeveelheid preventieve interventies, neemt de opbrengst-kostenratio over een periode van vijf jaar toe van €1,30 naar €1,32. Deze resultaten laten implementatiekosten buiten beschouwing.

Het toevoegen van preventie leidt tot een toename in het zorgbudget van 7%. Deze toename is relatief laag aangezien succesvolle preventie leidt tot kostenbesparing als gevolg van minder benodigde behandeling. Wanneer de inzet van interventies gericht op behandeling wordt verminderd om het totale zorgbudget gelijk te houden, wordt er onder gelijk budget meer gezondheidswinst behaald door het aanbieden van relatief meer kosteneffectieve preventie. Dit resulteert weer in een toename van de opbrengst-kostenratio van €1,30 tot €1,32, maar dit keer met een gelijkblijvend budget. Omdat budgets ten tijde van een economische crisis niet snel zullen toenemen, is het belangrijk om andere benaderingen te verkennen waarmee de kosteneffectiviteit van de zorg kan worden verbeterd.

Na vijf jaar preventie is een nieuw evenwicht nog niet bereikt van het aantal patiënten dat behandeling nodig heeft. Preventie vermindert de benodigde hoeveelheid behandeling, dus zo lang een nieuw epidemiologisch evenwicht nog niet is bereikt, leidt elk additioneel jaar preventie tot een overall meer gunstige opbrengst-kostenratio.

Het Markov model in hoofdstuk 4 bevat updates ten opzichte van het model in de publicatie van Lokkerbol ea (2014a) in termen van de onderliggende epidemiologie. Deze is nu gebaseerd op Nemesis-2, waarin gebruik werd gemaakt van de CIDI/DSM-IV (De Graaf ea, 2012b) en niet langer op Nemesis-1, waarin gebruik werd gemaakt van CIDI/DISM-III-R. Een andere update in deze nieuwe versie van DepMod heeft betrekking op de evidentie rondom preventie (van Zoonen ea, 2014). De conclusie is echter hetzelfde: preventie is kosteneffectief.

Nieuwe benaderingen om kosteneffectiviteit te verbeteren

Het gezondheidseconomisch onderzoeksgebied is continu in ontwikkeling. Nieuwe inzichten leiden ertoe dat andere routes richting zorginnovatie kunnen worden bewandeld. In de hoofdstukken 5 en 6 worden twee nieuwe aanpakken gepresenteerd om de kosteneffectiviteit van zorgsystemen te verbeteren; aanpakken waarbij een andere route wordt bewandeld dan het toevoegen van meer kosteneffectieve interventies aan het zorgsysteem, maar waarbij kosteneffectiviteit bijvoorbeeld wordt verbeterd door gelijktijdig te investeren en de-investeren in interventies gericht op dezelfde doelgroep. Deze aanpakken onderscheiden zich daarnaast doordat ze zich niet richten op één specifieke, nieuwe interventie, maar op de huidige, bestaande interventiemix om te komen tot een lijst met opties om de kosteneffectiviteit van het zorgsysteem te verbeteren. Deze lijst met opties dient als input voor beleidsmakers met betrekking tot systeeminnovatie, en kan vervolgens beoordeeld worden vanuit normatieve aspecten zoals ethiek, acceptatie, gepastheid, haalbaarheid en de bewijskracht van de gebruikte evidentie.

Hoofdstuk 5

In hoofdstuk 5 werd het verbeterpotentieel van de eerstelijns geestelijke gezondheidszorg in Nederland onderzocht door systematisch de impact op het zorgsysteem in kaart te brengen van kleine veranderingen in het bereik, de therapietrouw, de effectiviteit of de kosten van elke doelgroep – interventiecombinatie. Elk van deze parameters speelt een rol in de uiteindelijke kosteneffectiviteit van zorg. Het scannen van deze doelgroep – interventie – parameter combinaties in hun potentie om de kosteneffectiviteit van zorg te

verbeteren, kan er aan bijdragen dat innovatie wordt gericht op die specifieke gebieden binnen de zorg waar dit de grootste verbetering bewerkstelligt. De mate waarin een interventie de kosteneffectiviteit van zorg kan verbeteren wordt in belangrijke mate bepaald door de absolute grootte van de doelgroep. Hierdoor zijn het met name de hoog-prevalente interventies zoals farmacotherapie die het grootste verbeterpotentieel laten zien na een verbetering in bereik, therapietrouw, effectiviteit of kosten.

Het verbeteren van de kosteneffectiviteit van zorg door het vergroten van het bereik van interventies kan worden geïnterpreteerd als het gebruik maken van bestaande kosteneffectieve interventies (oogsten). Het verbeteren van therapietrouw, effectiviteit of de interventiekosten kan worden gezien als investeringsstrategieën met als doel om het verbeterpotentieel in deze gebieden te benutten (zaaien). Omdat de voorgestelde verbeteringen hypothetisch van aard zijn, is het van belang om deze verbeteringen weer kritisch te beschouwen vanuit een breder perspectief.

Hoofdstuk 6
In hoofdstuk 6 werd onderzocht hoe zorgsubstitutie kan leiden tot kosteneffectievere zorg. Hiervoor werd geanalyseerd wat de impact is van gelijktijdig investeren in relatief kosteneffectieve interventies en de-investeren in relatief kostenineffectieve interventies. Een algoritme werd ontwikkeld waarmee alle interventieparen in een zorgsysteem systematisch worden beoordeeld op hun potentieel om een zorgsysteem te creëren dat minstens zoveel gezondheid genereert onder een gelijk of kleiner zorgbudget (een dominant zorgsysteem) door gelijktijdige investering en de-investering. Het algoritme laat zien dat er over het algemeen evenveel mogelijkheden zijn om een dominant zorgsysteem te creëren als het aantal verschillende interventieparen. Aangezien het aantal mogelijke interventieparen snel toeneemt wanneer het aantal interventies in een zorgsysteem toeneemt, kan dit algoritme beleidsmakers voorzien van een lange lijst met verbeteropties waarvoor het niet benodigd is om te investeren in nieuwe interventies.

Het algoritme werd toegepast op het zorgsysteem voor depressie in Nederland, waarbij gebruik werd gemaakt van het gezondheidseconomische (Markov) model uit hoofdstuk 4. Er werden 11 interventieparen gevonden met het potentieel om een efficiënter zorgsysteem te creëren. De toepassing van het algoritme beperkt zich niet tot specifieke (bijv. Markov) modellen of tot de geestelijke gezondheidszorg, maar is algemeen toepasbaar. Aangezien uitkomsten geen toename in budget vereisen, maar wel de-investering omvatten, kan dit er toe leiden

dat politieke debatten minder op financiën gericht zullen zijn en meer op haalbaarheid en ethiek.

Bredere perspectieven

Voor de economische evaluaties in dit proefschrift en voor economische evaluaties in het algemeen, geldt dat voorgestelde innovaties naar verwachting vaak leiden tot meer kosteneffectieve zorg. Dit betekent uiteraard niet dat deze innovaties daarom meteen geïmplementeerd moeten worden; gezondheidseconomisch modelleren moet niet gezien worden als een autopilot voor innovatie. De economische evaluaties in de hoofdstukken 2–6 benaderen de zorg vanuit het perspectief van kosteneffectiviteit. De mate waarin een innovatie gewenst is vanuit het maatschappelijke perspectief hangt daarnaast af van andere factoren. Dit betekent dat de resultaten van gezondheidseconomische modellering idealiter gevolgd worden door een filteringproces waarin de innovaties worden beoordeeld op de mate waarin deze gepast, acceptabel en haalbaar zijn en gelijke toegang tot zorg verschaffen, alsook de sterkte van de bewijskracht. Dit normatieve filteringproces is gewenst bij de meeste gezondheidseconomische evaluaties, maar dit geldt nog meer voor de algoritmen in hoofdstukken 5 en 6, waar veel potentiële verbeteropties bekeken kunnen worden in dit filteringproces. Om zo goed mogelijk aan te sluiten bij criteria naast kosteneffectiviteit, wordt het proces van het ontwikkelen van een gezondheidseconomisch model idealiter ondersteund met input vanuit het klinische, patiënt- en beleidsmakerperspectief. Een goed voorbeeld hiervan is de ontwikkeling van multidisciplinaire richtlijnen, waar clinici en patiënten actief bij kunnen dragen aan de ontwikkeling van het conceptuele raamwerk van een gezondheidseconomisch model.

References

References

American Psychiatric Association. Diagnostic and Statistical Manual of Mental Disorders-III-Revised: DSM-III-R. Washington D.C., American Psychiatric Association, 1987.

Andersson G, Cuijpers P (2009). Internet-based and other computerized psychological treatments for adult depression: a meta-analysis. Cogn Behav Ther 38: 196-205.

Andrews G, Sanderson K, Beard J (1998). Burden of disease: methods of calculating disability from mental disorder. Brit J Psychiatry 173: 123–131.

Andrews G, Issakidis C, Sanderson K, Corry J, Lapsley H (2004). Utilising survey data to inform public policy: comparison of the cost-effectiveness of treatment of ten mental disorders. Brit J Psychiatry 184: 526-533.

Annemans L, Brignone M, Druais S, De Pauw A, Gauthier A, et al. (2014). Cost-effectiveness analysis of pharmaceutical treatment options in the first-line management of major depressive disorder in Belgium. PharmacoEconomics 32(5): 479-493.

Aragonèsa E, López-Cortacansa G, Sánchez-Irisoc E, Piñola JL, Caballeroa A, et al. (2014). Cost-effectiveness analysis of a collaborative care programme for depression in primary care. J Affect Disord 159: 85–93.

Arroll B, MacGillivray S, Ogston S, Reid I, Sullivan F, et al. (2005). Efficacy and tolerability of tricyclic antidepressants and SSRIs compared with placebo for treatment of depression in primary care: a meta-analysis. Ann Fam Med 3: 449-456.

Babor TF, Higgins-Biddle JC, Saunders JB, Monteiro MG (2001). The alcohol use disorders identification test: guidelines for use in primary care. 2nd edition. Geneva, Switzerland: World Health Organization.

Bakker IM, Terluin B, Van Marwijk HWJ, van der Windt DAWM, Rijmen F, et al. (2007). A cluster-randomized trial evaluating an intervention for patients with stress related mental disorders and sick leave in primary care. PLoSClinicalTrials.

Barbosa C, Taylor B, Godfrey C, Rehm J, Parrott S, et al. (2010). Modelling lifetime QALYs and health care costs from different drinking patterns over time: a Markov model. Int J Methods Psychiatr Res 19(2): 97-109.

Barrett B, Byford S (2003). Acceptable mental health care? A new decision-making tool for cost-effectiveness analysis. Mental Health Research Review 9: 39-41.

Beaglehole R, Bonita R (2009). Alcohol: a global health priority. Lancet 373(9682): 2173-2174.

Begg C, Cho M, Eastwood S, Horton R, Moher D, et al. (1996). Improving the quality of reporting of randomized controlled trials. The CONSORT statement. JAMA 276: 637-639.

Begg SJ, Vos T, Barker B, Stanley L, Lopez AD (2008). Burden of disease and injury in Australia in the new millennium: measuring health loss from diseases, injuries and risk factors. Med J Australia 188: 36–40.

Beich A, Thorsen T, Rollnick S (2003). Screening in brief intervention trials targeting excessive drinkers in general practice: systematic review and meta-analysis. Br Med J 327: 536-542.

Benegal V, Chand PK, Obot IS (2009). Packages of care for alcohol use disorders in low- and middle-income countries. PLoS Med 6(10): e1000170.

Berghmans R, Berg M, van den Burg M, ter Meulen R (2004). Ethical issues of cost effectiveness analysis and guideline setting in mental health care. J Med Ethics 30: 146-150.

Berto P, D'Ilario D, Ruffo P, Virgilio RD, Rizzo F (2000). Depression: cost-of-illness studies in the international literature: a review. J Ment Health Policy Econ 3:3-10.

Bijl RV, Van Zessen G, Ravelli A (1998). The Netherlands Mental Health Survey and Incidence Study (NEMESIS): objectives and design. Soc Psychiatry Psychiatr Epidemiol 33: 581–586.

Bijl RV, de Graaf R, Ravelli A, Smit F, Vollebergh WAM (2002). Gender and age-specific first incidence of DSM-III-R psychiatric disorders in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Soc Psychiatry Psychiatr Epidemiol 37: 372 – 379.

Birch S, Donaldson C (2003). Valuing the benefits and costs of health care programmes: where's the 'extra' in extra-welfarism? Soc Sci Med 56(5): 1121–1133.

Blankers M, Koeter MW, Schippers GM (2011). Internet therapy versus internet self-help versus no treatment for problematic alcohol use: A randomized controlled trial. J Consult Clin Psychol 79(3): 330-341.

Blankers M, Nabitz U, Smit F, Koeter MW, Schippers GM (2012). Economic evaluation of internet-based interventions for harmful alcohol use alongside a pragmatic randomized controlled trial. J Med Internet Res 14(5): e134.

Boon B, Risselada A, Huiberts A, Riper H, Smit F (2011). Curbing alcohol use in male adults through computer generated personalized advice: randomized controlled trial. J Med Internet Res 13(2): e43.

Botsis T, Hartvigsen G (2008). Current status and future perspectives in telecare for elderly people suffering from chronic diseases. J Telemed Telecare 14: 195-203.

Bowie C, Beck S, Bevan G, Raftery J, Silverton F, et al. (1997). Estimating the burden of disease in an English region. J Public Health Med 19: 87–92.

Bowman D, Scogin F, Lyrene B (1995). The efficacy of self-examination therapy and cognitive bibliotherapy in the treatment of mild to moderate depression. Psychotherapy Research 5: 131-140.

Bowman D, Scogin F, Floyd M (1997). Efficacy of self-examination therapy in the treatment of generalized anxiety disorder. J Couns Psychol 44: 267-273.

Brazier J, Roberts J, Deverill M (2002). The estimation of a preference-based measure of health from the SF-36. J Health Econ 21: 271–292.

Briggs AH, Weinstein MC, Fenwick EAL, Karnon J, Sculpher MJ, et al. (2012). Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–6. Med Decis Making, 32(5): 722–732.

Briggs A, Claxton K, Sculpher MJ (2006). Decision modelling for health economic evaluation. Oxford, England, Oxford University Press.

Brouwer WBF, Koopmanschap MA (2000). On the economic foundations of CEA. Ladies and gentlemen, take your positions! Health Econ 19: 439–459.

Byers AL, Covinsky KE, Barnes DE, Yaffe K (2012). Dysthymia and depression increase risk of dementia and mortality among older veterans. Am J Geriatr Psychiatry 20: 664-672.

Caro JJ, Briggs AH, Siebert U, Kuntz KM (2012). Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–1. Med Decis Making 32(5): 667–677.

Carter R (2001). The Macro Economic Evaluation Model (MEEM): An approach to priority setting in the health sector, PhD thesis. Faculty of Business and Economics. Melbourne, Monash University.

Carter R, Vos T, Moodie M, Haby M, Magnus A, et al. (2008). Priority setting in health: origins, description and application of the Australian Assessing Cost Effectiveness (ACE) initiative. Expert Rev Pharmacoecon Outcomes 8(6): 593-617.

Chandra A, Skinner J (2012). Technology growth and expenditure growth in health care. J Econ Lit 50(3): 645–680.

Chinn S (2000). A simple method for converting an odds ratio to effect size for use in meta-analysis. Stat Med 19(22): 3127-3131.

Chisholm D, Sanderson K, Ayoso-Mateos JL, Saxena S (2004a). Reducing the global burden of depression: population-level analysis of intervention cost-effectiveness in 14 world regions. Br J Psychiatry 184: 393-403.

Chisholm D, Rehm J, Van Ommeren M, Monteiro M (2004b). Reducing the global burden of hazardous alcohol use: comparative cost-effectiveness analysis. J Stud Alcohol 65: 782–793.

Chisholm D (2005). Choosing cost-effective interventions in psychiatry: results from the CHOICE programme of the World Health Organization. World Psychiatry 4(1): 37-44.

Chisholm D (2007). Mental health system financing in developing countries: policy questions and research responses. Epidemiol Psychiatr Soc 16(4): 282-288.

Chisholm D, Gureje O, Saldivia S, Villalón Calderón M, Wickremasinghe R, et al. (2008). Schizophrenia treatment in the developing world: an interregional and multinational cost-effectiveness analysis. Bull World Health Organ 86(7): 542-51.

Chisholm D, Stanciole AE, Tan-Torres Edejer T, Evans DB (2010). Economic impact of disease and injury: counting what matters. BMJ 340: c924.

Christensen H, Griffiths KM, Farrer L (2009). Adherence in internet interventions for anxiety and depression. J Med Internet Res 11(2): e13.

Council for Public Health and Health Care (2006). Sensible and sustainable care. Zoetermeer, Netherlands: Council for Public Health and Health Care (RVZ).

Cuijpers P (1998). A psychoeducational approach to the treatment of depression: a meta-analysis of Lewinsohn's "Coping with Depression" course. Behavior Therapy 29: 521-533.

Cuijpers P, Smit F (2002). Excess mortality in depression: a meta-analysis of community studies. J Affect Dis, 72: 227-236.

Cuijpers P, Van Straten A, Warmerdam AM, Smits N (2007). Characteristics of effective psychological treatments of depression: a metaregression analysis. Psychotherapy Research 18(2): 225-236.

Cuijpers P, van Straten A, Smit F, Mihalopoulos C, Beekman A (2008a). Preventing the onset of depressive disorders: A meta-analytic review of psychological interventions. Am J Psychiatry, 165: 1272-1280.

Cuijpers P, van Straten A, Andersson G (2008b). Internet-administered cognitive behavior therapy for health problems: a systematic review. J Behav Med 31: 169-177.

Cuijpers P, Straten A, Andersson G, van Oppen P (2008c). Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J Consult Clin Psychol 76(6): 909-922.

Cuijpers P, Van Straten A, Warmerdam MA, Andersson G (2009). Psychotherapy versus combination of psychotherapy and pharmacotherapy in the treatment of depression: a meta-analysis. Depress Anxiety 26: 279-288.

Cuijpers P, Donker T, van Straten A, Li J, Andersson G (2010). Is guided self-help as effective as face-to-face psychotherapy for depression and anxiety disorders? A systematic review and meta-analysis of comparative outcome studies. Psychol Med 40(12): 1943-1957.

Cuijpers P, Beekman ATF, Reynolds CF (2012). Preventing depression, a global priority. JAMA 307(10): 1033–1034.

Craig BM, Busschbach JJ, Salomon JA (2009). Keep it simple: ranking health states yields values similar to cardinal measurement approaches. J Clin Epidemiol 62: 296–305.

Culyer AJ (1990). Commodities, characteristics of commodities, characteristics of people, utilities and the quality of life. In: Baldwin S, Godfrey C, Propper C (Eds.), The Quality of Life: Perspectives and Policies. Routledge, London.

Davidson T (2009). How to include relatives and productivity loss in a cost-effectiveness analysis. Dissertation, Linköping University, Sweden.

Davis WT, Campbell L, Tax J, Lieber CS (2002). A trial of "standard" outpatient alcoholism treatment vs. a minimal treatment control. J Subst Abuse Treat 23(1): 9-19.

De Graaf R, Bijl RV, Spijker J, Beekman ATF, Vollebergh WAM (2003). Temporal sequencing of lifetime mood disorders in relation to comorbid anxiety and substance disorders. Soc Psychiatry Psychiatr Epidemiol 38: 1–11.

De Graaf R, Tuithof M, van Dorsselaer S, ten Have M (2012a). Comparing the effects on work performance of mental and physical disorders. Soc Psychiatry Psychiatr Epidemiol 47(11): 1873-1883.

De Graaf R, ten Have M, Tuithof M, van Dorsselaer S (2012b). Incidentie van psychische aandoeningen: Opzet en eerste resultaten van de tweede meting van de studie NEMESIS-2. Trimbos Institute, Utrecht.

De Mast J, Does RJMM, De Koning H, Lokkerbol J (2012). Lean Six Sigma for Services and Healthcare. Beaumont Quality Productions, Alphen aan den Rijn.

Donker T, van Straten A, Riper H, Marks I, Andersson G, et al. (2009). Implementation of internet-based preventive interventions for depression and anxiety: role of support? The design of a randomized controlled trial. Trials 10: 59.

Drost RMWA, Paulus ATG, Ruwaard D, Evers SMAA (2014). Handleiding intersectorale kosten en baten van (preventieve) interventies: classificatie, identificatie en kostprijzen. Maastricht University, Faculty of Health, Medicine and Life Sciences CAPHRI, School for Public Health and Primary Care, Department of Health Services Research, ISBN: 9789461593429.

Drummond MF (1981). Welfare economics and cost benefit analysis in health care. Scot J Polit Econ 28: 125-145.

Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL (2005). Methods for the economic evaluation of health care programmes. Oxford, England, Oxford University Press.

Drummond M, Barbieri M, Cook J, Glick HA, Lis J, et al. (2009). Transferability of economic evaluations across jurisdictions: ISPOR Good Research Practices Task Force Report. Value Health 12(4): 409 – 418.

Eddama O, Coast J (2008). A systematic review of the use of economic evaluation in local decision-making. Health Policy 86: 129-141.

Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, et al. (2012). Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Med Decis Making 32(5): 733–743.

Ekers D, Richards D, Gilbody S (2008). A meta-analysis of randomised trials of behavioural treatment of depression. Psychol Med 38: 611-623.

Elshaug AG, Hiller JE, Moss JR (2008). Exploring policy-makers' perspectives on disinvestment from ineffective healthcare practices. Int J Technol Assess 24(1): 1-9.

ESEMeD (2004). Disability and quality of life impact of mental disorders in Europe: results from the European Study of the Epidemiology of Mental Disorders (ESEMeD) project. Acta Psychiatr Scand 109 (suppl): 38–46.

EuroQoL Group (1990). EuroQoL: a new facility for the measurement of health-related quality of life. Health Policy 16: 199–208.

Ezzati M, Lopez AD, Rodgers A, Murray C (2004). Comparative quantification of health risks: global and regional burden of diseases attributable to selected major risks. Geneva, Switzerland: World Health Organization.

Fortney JC, Pyne JM, Edlund MJ, Williams DK, Robinson DE, et al. (2007). A randomized trial of telemedicine-based collaborative care for depression. J Gen Intern Med 22: 1086-1093.

Fournier JC, DeRubeis RJ, Hollon SD, Dimidjian S, Amsterdam JD, et al. (2010). Antidepressant drug effects and depression severity: a patient-level meta-analysis. J Am Med Assoc 303(1): 47-53.

Gafni A (1994). The standard gamble method: what is being measured and how it is interpreted. Health Serv Res 29: 207–224.

Garner S, Docherty M, Sommer J, Sharma T, Choudhury M, et al. (2013). Reducing ineffective practice: challenges in identifying low-value health care using Cochrane systematic reviews. J Health Serv Res Policy 18(1): 6-12.

Gega L, Kenwright M, Mataix-Cols D, Cameron R, Marks IM (2005). Screening people with anxiety/depression for suitability for guided self-help. Cognitive Behaviour Therapy 34: 16-21.

GGZ guidelines (2009). Multidisciplinary guideline for disorders in the use of alcohol. Utrecht, Trimbos Institute.

GGZ guidelines (2011a). Multidisciplinary guideline for anxiety disorders, second revision. Utrecht, Trimbos Institute.

GGZ guidelines (2011b). Multidisciplinary guideline for depression, second revision. Utrecht, Trimbos Institute.

Gmel G, Gutjahr E, Rehm J (2003). How stable is the risk curve between alcohol and all-cause mortality and what factors influence the shape? A precision-weighted hierarchical meta-analysis. Eur J Epidemiol 18(7): 613–642.

Greenberg PE, Birnbaum HG (2005). The economic burden of depression in the US: societal and patient perspectives. Exp Opin Pharmacother 6: 369-376.

Hakkaart-van Roijen L, Tan SS, Bouwmans CAM (2010). Handleiding voor kostenonderzoek. Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. Instituut voor Medical Technology Assessment, Erasmus Universiteit Rotterdam.

Hakkaart-van Roijen L (2002). Manual Trimbos/iMTA Questionnaire for costs associated with psychiatric illness (in Dutch). Rotterdam: Institute for Medical Technology Assessment.

Harrison V, Proudfoot J, Wee PP, Parker G, Pavlovic DH, et al. (2011). Mobile mental health: review of the emerging field and proof of concept study. JMH 20(6): 509-524.

Haug T, Nordgreen T, Ost LG, Havik OE (2012). Self-help treatment of anxiety disorders: A meta-analysis and meta-regression of effects and potential moderators. Clinical psychology review 32: 425-445.

Haverman M, Boon B, Riper H (2009). Therapietrouw & onderzoeksuitval bij e-mental health interventies. First international E-Mental Health Summit.

Hawthorne G, Richardson J, Day NA (2001). A comparison of the Assessment of Quality of Life (AQoL) with four other generic utility instruments. Ann Med 33: 358-370.

Higgins CB, Roeske WR, Karliner JS, O'Rourke R, Berk RN (1976). Predictive factors and mechanism of arrhythmias and myocardial ischaemic changes in elderly patients during barium enema. Br J Radiol 49(588): 1023-1027.

Hoffman D, Dukes EM, Wittchen HU (2008). Human and economic burden of Generalized Anxiety Disorder. Depress Anxiety 25(1): 72-90.

Holman AJ, Serfaty MA, Leurent BE, King MB (2011). Cost-effectiveness of cognitive behavior therapy versus talking and usual care for depressed older people in primary care. BMC Health Serv Res 11: 33.

Hoomans T, Evers SMAA, Ament AJHA, Hubben MWA, van der Weijden T, et al. (2007). The methodological quality of economic evaluations of guideline implementation into clinical practice: a systematic review of empiric studies. Value Health 10(4): 305-316.

Hox JJ (2010). Multilevel analysis: techniques and applications. Taylor and Francis Group, New York.

Hurley J (1998). Welfarism, extra-welfarism and evaluative economic analysis in the health care sector. In: Barer ML, Getzen TE, Stoddard GL (Eds.), Health, Health Care and Health Economics: Perspectives on Distribution. John Wiley & Sons Ltd., Chichester.

Hurley J (2000). An overview of the normative economics of the health care sector. In: Culyer AJ, Newhouse J (Eds.), Handbook of Health Economics. Elsevier Science, Amsterdam.

Husereau D, Drummond M, Petrou S, Carswell C, Moher D, et al. (2013). Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMC Med 11: 80.

Hutubessy R, Chisholm D, Edejer TTT (2003). Generalized cost-effectiveness analysis for national-level priority-setting in the health sector. Cost Eff Resour Alloc 1:8.

Hsiao W, Heller PS (2007). What should macroeconomists know about health care policy? IMF Working Paper, WP/07/13.

IBM Corp. (2013). IBM SPSS Statistics for Windows, Version 22.0. Armonk, NY: IBM Corp.

Issakidis C, Sanderson K, Corry J, Andrews G, Lapsley H (2004). Modelling the population cost-effectiveness of current and evidence based optimal treatment for anxiety disorders. Psychol Med 34: 1-17.

Johns B, Baltussen R, Hutubessy R (2003). Programme costs in the economic evaluation of health interventions. Cost Eff Resour Alloc 1: 1.

Karanikolos M, Mladovsky P, Cylus J, Thomson S, Basu S, et al. (2013). Financial crisis, austerity and health in Europe. Lancet 381(9874): 1323-1331.

Karnon J, Stahl J, Alan B, Caro JJ, Mar J, et al. (2012). Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–4. Med Decis Making 32(5): 701–711.

Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, et al. (2008). Initial severity and antidepressant benefits: a meta-analysis of data submitted to the food and drug administration. Plos Med 5(2): e45.

Konnopka A, König H–H (2009). The health and economic consequences of moderate alcohol consumption in Germany 2002. Value Health 2: 253–261.

Kosten TR, O'Connor PG (2003). Management of drug and alcohol withdrawal. N Engl J Med 348(18): 1786-1795.

Kraemer KL, Roberts MS, Horton NJ, Palfai T, Samet JH, et al. (2005). Health utility ratings for a spectrum of alcohol-related health states. Med Care 43(6): 541-550.

Lamers LM, Stalmeier PFM, McDonnell J, Krabbe PFM, Van Busschbach JJ (2005). Kwaliteit van leven meten in economische evaluaties: het Nederlands EQ-5D-tarief. Ned Tijdschr Geneesk 149: 1574-1578.

Lancet Global Mental Health Group (2007). Scale up services for mental health: a call for action. Lancet 370: 1241–1252.

Landelijk Kwaliteitsinstituut CBO & Trimbos Instituut (2009). Multidisplinaire richtlijn stoornissen in het gebruik van alcohol Multidisciplinary Guideline Alcohol Use Disorders. Utrecht, the Netherlands: Trimbos Institute.

Lauer JA, Röhrich K, Wirth H, Charette C, Gribble S, et al. (2003). PopMod: a longitudinal population model with two interacting disease states. Cost Eff Res Alloc 1: 6.

Le Lay A, Despiegel N, Francois C, Duru G (2006). Can discrete event simulation be of use in modelling major depression? Cost Eff Res Alloc 4: 19.

Licht-Strunk E, van der Windt DA, van Marwijk HW, de Haan M, Beekman AT (2007). The prognosis of depression in older patients in general practice and the community. A systematic review. Fam Pract 24: 168-180.

Little RJA and Rubin DB (1987). Statistical analysis with missing data, New York: John Wiley.

Lokkerbol J, Schotman MAM, Does RJMM (2011). Quality Quandaries: Personal injuries: a case study. Quality Engineering 24(1): 102-106.

Lokkerbol J, Does RJJM, de Mast J, Schoonhoven S (2012a). Improving processes in financial service organizations: where to begin? Int J Quality Reliability Management 29 (9): 981 – 999.

Lokkerbol J, Molenaar M, Does RJMM (2012b). Quality Quandaries: An efficient public sector. Quality Engineering 24: 431–435.

Lokkerbol J, Adema D, de Graaf R, ten Have M, Cuijpers P, et al. (2013). Non-fatal burden of disease due to mental disorders in the Netherlands. Soc Psychiatry Psychiatr Epidemiol 48(10): 1591-1599.

Lokkerbol J, Adema D, Cuijpers P, Reynolds CF, Schulz R, et al. (2014a). Improving the cost-effectiveness of a health care system for depressive disorders by implementing telemedicine: a health economic modelling study. Am J Geriatr Psychiatry 22(3): 253-262.

Lokkerbol J, Weehuizen R, Mavranezouli I, Mihalopoulos C, Smit F (2014b). Mental health care system optimization from a health-economics perspective: where to sow and where to reap? J Ment Health Policy Econ 17: 51-60.

Mann K (2004). Pharmacotherapy of alcohol dependence: a review of the clinical data. CNS Drugs 18(8): 485-504.

Mathers CD, Loncar D (2006). Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med 3(11): e442.

Mathers CD, Vos T, Stevenson E, Begg SJ (2001). The burden of disease and injury in Australia. Bull World Health Organ 79: 1076–1084.

Mauskopf JA, Sullivan SD, Annemans L, Caro J, Mullins CD, et al. (2007). Principles of good practice for budget impact analysis: report of the ISPOR Task Force on Good Research Practices— Budget Impact Analysis. Value Health 10(5): 336-347.

Meijer S, Smit F, Schoemaker C, Cuijpers P (2006). Gezond verstand: evidence-based preventie van psychische stoornissen Common sense: evidence-based prevention of mental disorders De Bilt: RIVM, centrum Volksgezondheid Toekomst Verkenningen en Trimbos-instituut, centrum Preventie en Kortdurende Interventie. In: Gezond verstand: evidence-based preventie van psychische stoornissen. Bilthoven: RIVM.

Menken M, Munsat TL, Tolle JF (2000). The global burden of disease study: implications for neurology. Arch Neurol 57: 418– 420.

Meulenbeek P, Willemse G, Smit F, van Balkom A, Spinhoven P, et al. (2010). Early intervention in panic: pragmatic randomized controlled trial. Br J Psychiatry 196: 326-331.

Mihalopoulos C (2011). The cost-effectiveness of preventive interventions for mental disorders: the application and evaluation of a priority-setting model, PhD Thesis. School of Population Health, The University of Melbourne.

Mihalopoulos C, Vos T, Pirkis J, Smit F, Carter R (2011a). Do indicated preventive interventions represent good value-for-money? Aust N Z J Psychiatry 45: 36-44.

Mihalopoulos C, Vos T, Pirkis J, Carter R (2011b). The economic analysis of prevention in mental health programs. Annu Rev Clin Psychol 7: 169-201.

Mihalopoulos C, Vos T, Pirkis J, Carter R (2012). The population cost-effectiveness of interventions designed to prevent childhood depression. Pediatrics 129(3): e723-30.

Mihalopoulos C, Vos T (2013a). Cost-effectiveness of preventive interventions for depressive disorders: an overview. Pharmacoecon Outcomes Res 13(2): 237-42.

Mihalopoulos C, Carter R, Pirkis J, Vos T (2013b). Priority-setting for mental health services. J Ment Health 22(2): 122-134.

Mohapatra S, Patra J, Popova S, Duhig A, Rehm J (2010). Social cost of heavy drinking and alcohol dependence in high-income countries. Int J Public Health 55(3): 149-157.

Moher D, Schulz KF, Altman DG (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomised trials. Lancet 357: 1191-1194.

Molen HT Van der, Kluytmans F, Kramer M (1995). Gespreksvoering. Vaardigheden en modellen Groningen: Open Universiteit Heerlen, the Netherlands.

Moyer A, Finney JW, Swearingen CE, Vergun P (2002). Brief interventions for alcohol problems: a meta-analytic review of controlled investigations in treatment-seeking and non-treatment-seeking populations. Addiction 97(3): 279-292.

Muñoz RF, Cuijpers P, Smit F, Barrera AZ, Leykin Y (2010). Prevention of major depression. Ann Rev Clin Psychol 6: 181-212.

Murray CJL, Lopez AD (1996). The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020. Vol 1, Mass: Harvard School of Public Health, on behalf of the World Health Organization and the World Bank, Cambridge.

Murray C, Lopez A (1997). Global mortality, disability, and the contribution of risk factors: Global Burden of Disease Study. Lancet 349: 1436-1442.

Murray CJL, Evans DB, Acharya A, Baltussen RMPM (2000). Development of WHO guidelines on generalized cost-effectiveness analysis. Health Econ 9: 235-251.

Murray C, Lopez A, Mathers CD, Stein C (2001). The Global Burden of Disease 2000 Project: aims, methods and data sources. Geneva, Switzerland: World Health Organization.

Murray CJ, Lopez AD, Wibulpolprasert S (2004). Monitoring global health: time for new solutions. BMJ 329(7474): 1096-1100.

Murray CJ, Frenk J (2008). Health metrics and evaluation: strengthening the science. Lancet 371(9619): 1191-1199.

Murray CJ, Vos T, Lozano R (2012). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380(9859): 2197-2223.

Murray CJ, Richards MA, Newton JN, Fenton KA, Anderson HR, et al. (2013a). UK health performance: findings of the Global Burden of Disease Study 2010. Lancet 381(9871): 997-1020.

Murray CJ, Lopez AD (2013b). Measuring the global burden of disease. N Engl J Med 369(5): 448-457.

National Institute of Mental Health (NIMH): Treatment for adolescents with depression study (TADS) pharmacotherapy treatment manual.

NICE Guidelines (2011). Generalised anxiety disorder in adults: management in primary, secondary and community care. National Clinical Guideline Number 113.

NICE Guidelines (2014). Psychosis and Schizophrenia in adults. The NICE guideline on treatment and management – updated edition 2014. National Clinical Guideline Number 178.

Niessen LW, Bridges J, Lau BD, Wilson RF, Sharma R, et al. (2012). Assessing the impact of economic evidence on policymakers in health care—a systematic review. Agency for Healthcare Research and Quality. Rockville, Maryland.

Nordin S, Carlbring P, Cuijpers P, Andersson G (2010). Expanding the limits of bibliotherapy for panic disorder: randomized trial of self-help without support but with a clear deadline. Behavior Therapy 41(3): 267-276.

OECD Stats. Available from http://stats.oecd.org/Index.aspx?DataSetCode=SHA, accessed 05-02-2014.

OECD STATS. PPPs and exchange rates. http://stats.oecd.org/Index.aspx?DataSetCode1⁄4SNA_TABLE4. Accessed June 1, 2014.

Oostenbrink JB, Bouwmans C, Koopmanschap MA (2004). Manual for costing: methods and standard costs for economic evaluations in health care (in Dutch). Diemen, the Netherlands: Health Insurance Board.

Pearson S, Littlejohns P (2007). Reallocating resources: how should the national institute for health and clinical excellence guide disinvestment efforts in the National Health Service? J Health Serv Res Policy 12(3): 160-165.

Pillai A, Nayak MB, Greenfield TK, Bond JC, Nadkarni A, et al. (2012). Patterns of alcohol use, their correlates, and impact in male drinkers: a population-based survey from Goa India. Soc Psychiatry Psychiatr Epidemiol 48(2): 275–282.

Pitman R, Fisman D, Zaric GS, Postma M, Kretzschmar M, et al. (2012). Dynamic transmission modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group–5. Med Decis Making 32(5): 712–721.

Polisena J, Clifford T, Elshaug AG, Mitton C, Russell E, et al. (2013). Case studies that illustrate disinvestment and resource allocation decision-making processes in health care: a systematic review. Int J Technol Assess 29(2): 174-184.

Proctor EK, Landsverk J, Aarons G, Chambers D, Glisson C, et al. (2009). Implementation research in mental health services: an emerging science with conceptual, methodological, and training challenges. Adm Policy Ment Health 36: 24-34.

Ramsey S, Willke R, Briggs A, Brown R, Buxton M, et al. (2005). Good research practice for cost-effectiveness analysis alongside clinical trial: the ISPOR RCT-CEA Task Force report. Value Health 8(5): 521-533.

Reid MC, Fiellin DA, O'Connor PG (1999). Hazardous and harmful alcohol consumption in primary care. Arch Intern Med 159(15): 1681-1689.

Rehm J, Room R, Monteiro M, Gmel G, Graham K, et al. (2003). Alcohol as a risk factor for global burden of disease. Eur Addict Res 9: 157–164.

Rehm J, Taylor B, Room R (2006). Global burden of disease from alcohol, illicit drugs and tobacco. Drug Alcohol Rev 25(6): 503-513.

Rehm J, Mathers C, Popova S, Thavorncharoensap M, Teerawattananon Y, et al. (2009). Global burden of disease and injury and economic cost attributable to alcohol use and alcohol-use disorders. Lancet 373: 2223-2233.

Richards D, Richardson T (2012). Computer-based psychological treatment for depression: a systematic review and meta-analysis. Clin Psychol Rev 32(4): 329-342.

Riper H, Smit F, Van der Zanden R, Conijn B, Kramer J, et al. (2007). E-mental health: high tech, high touch, high trust. Utrecht, The Netherlands, Trimbos Institute.

Riper H, Kramer J, Smit F, Conijn B, Schippers G, et al. (2008). Web-based self-help for problem drinkers: a pragmatic randomized trial. Addict 103(2): 218-227.

Riper H, Andersson G, Christensen H, Cuijpers P, Lange A, et al. (2010). Theme issue on E-mental health: a growing field in internet research. J Med Internet Res 12(5): e74.

Riper H, Spek V, Boon B, Conijn B, Kramer J, et al. (2011). Effectiveness of e-self-help interventions for curbing adult problem drinking: a meta-analysis. J Med Internet Res 13(2): e42.

Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, et al. (2012). Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–2. Med Decis Making 32(5): 678–689.

Rogers CR (1951). The attitude and orientation of the counsellor. In client-centered therapy, its current practice, implications, and theory Edited by: Carmichael L. Boston: Houghton Mifflin: 19-64.

Room R, Babor T, Rehm J (2005). Alcohol and public health. Lancet 365(9458): 519-530.

Ruwaard J, Schrieken B, Schrijver M, Broeksteeg J, Dekker J, et al. (2009). Standardized web-based cognitive behavioural therapy of mild to moderate depression: a randomized controlled trial with a long-term follow-up. Cognitive Behaviour Therapy 38: 206-221.

Saarni SI, Ari JS, Sintonen H, Pirkola S, Koskinen S, et al. (2007). Impact of psychiatric disorders on health-related quality of life: general population survey. Brit J Psychiatry 190: 326–332.

Salomon JA, Vos T, Hogan DR, Gagnon M Naghavi M (2012). Common values in assessing health outcomes from disease and injury: disability weights measurement study for the Global Burden of Disease Study 2010. Lancet 380: 2129-43.

Sanderson K, Andrews G, Corry J, Lapsley H (2004). Using the effect size to model change in preference values from descriptive health status. Qual Life Res 13(7): 1255-1264.

Schulz KF, Altman DG, Moher D, the CONSORT group (2010). CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials. Trials 11: 32.

Sendi P, Gafni A, Birch S (2002). Opportunity costs and uncertainty in the economic evaluation of health care interventions. Health Econ 11: 23-31.

Shewhart WA (1939). Statistical method from the viewpoint of quality of control. Dover Publications.

Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, et al. (2012). State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force–3. Med Decis Making 32(5): 690–700.

Sindern J, Schröder-Bernhardi D, Slawik L (2010). Secondary data for budget impact analyses. Gesundheitswesen 72(6): 356-362.

Smit F, Cuijpers P, Oostenbrink J, Batelaan N, De Graaf R, et al. (2006a). Costs of nine common mental disorders: implications for curative and preventive psychiatry. J Ment Health Policy Econ 9: 193–200.

Smit F, Willemse G, Koopmanschap M, Onrust S, Cuijpers P, et al. (2006b). Cost-effectiveness of preventing depression in primary care patients. Brit J Psychiatry 188: 330-336.

Smit F, Cuijpers P, Oostenbrink J, Batelaan N, de Graaf R, et al. (2006c). Excess costs of common mental disorders: population-based cohort study. J Ment Health Policy Econ 9: 193-200.

Smit F, Van Laar M, Croes E, Busschbach J (2008). Ziektelastgewichten voor misbruik, schadelijk gebruik en afhankelijkheid van alcohol en drugs. Disability weights for abuse, harmful use and dependence of alcohol and drugs. Utrecht, the Netherlands: Trimbos Institute.

Smit F, Lokkerbol J, Riper H, Majo CM, Boon B, et al. (2011). Modelling the cost-effectiveness of health care systems for alcohol use disorders: how implementation of eHealth interventions improve cost-effectiveness. J Med Internet Res 13(3): e56.

Spek V, Cuijpers P, Nylicek I, Riper H, Keyzer J, et al. (2007a). Internet-based cognitive behaviour therapy for symptoms of depression and anxiety: a meta-analysis. Psychol Med 37: 319-328.

Spek V, Nyklicek I, Smits N, Cuijpers P, Riper H, et al. (2007b). Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. Psychol Med 37: 1797-1806.

Spijker J, De Graaf R, Van Bijl R, Beekman A, Ormel J, et al. (2002). Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry 181: 208-213.

Stouthard M, Essink-Bot M-L, Bonsel G, Barendregt J, Kramers P, et al. (1997). Disability weights for diseases in the Netherlands. Department of Public Health, Erasmus University, Rotterdam.

Sugden R, Williams A (1978). The principles of practical cost-benefit analysis. Oxford University Press.

Tan-Torres Edejer T, Baltussen R, Adam T, Hutubessy R, Acharya A, et al. (2003). WHO guide to cost-effectiveness analysis. World Health Organization, Geneva.

Tate DF, Finkelstein EA, Khavjou O, Gustafson A (2009). Cost effectiveness of internet interventions: review and recommendations. Ann Behav Med 38(1): 40-45.

Ter Smitten MH, Smeets RMW, van den Brink W (1998). Composite International Diagnostic Interview (CIDI), Version 2.1, 12-months (in Dutch). World Health Organization, Amsterdam/Geneva.

Tonigan JS, Toscova R, Miller WR (1996). Meta-analysis of the literature on Alcoholics Anonymous: sample and study characteristics moderate findings. J Stud Alcohol 57(1): 65-72.

Tsourapas A, Frew E (2011). Evaluating 'success' in programme budgeting and marginal analysis: a literature review. J Health Serv Res Policy 16(3): 177-183.

United Nations (2013). World population prospects: the 2012 revision, volume I: comprehensive tables. Department of Economic and Social Affairs/Population Division.

Ustun TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJ (2004). Global burden of depressive disorders in the year 2000. Br J Psychiatry 184: 386-392.

Valmaggia LR, McCrone P, Knapp M, Woolley JB, Broome MR, et al. (2009). Economic impact of early intervention in people at high risk of psychosis. Psychol Med 39: 1617–1626.

van't Veer-Tazelaar PJ, van Marwijk HW (2011). Prevention of late-life anxiety and depression has sustained effects over 24 months: a pragmatic randomized trial. Am J Geriatr Psychiatry 19: 230-239.

Van Den berg M, Smit F, Vos T, van Baal PHM (2011). Cost-effectiveness of opportunistic screening and minimal contact psychotherapy to prevent depression in primary care patients. PloS One 6: e22884.

Van der Klink JJ, Blonk RW, Schene AH, van Dijk FJ (2001). The benefits of interventions for work-related stress. Am J Public Health 91(2): 270-276.

Van Dijck D, Knibbe RA (2005). De prevalentie van probleemdrinken in Nederland: een algemeen bevolkingsonderzoek. Universiteit Maastricht.

Van Hout BA, Al MJ, Gordon GS, Rutten FFH (1994). Costs, effects and C/E-ratios alongside a clinical trial. Health Econ 3: 309-319.

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

Van Straten A, Cuijpers P, Smits N (2008). Effectiveness of a web-based self-help intervention for symptoms of depression, anxiety, and stress: randomized controlled trial. J Med Internet Res 10: e7.

Van Zoonen K, Buntrock C, Ebert DD, Smit F, Reynolds CF, et al. (2014). Preventing the onset of major depressive disorder: A meta-analytic review of psychological interventions. Int J Epidemiol 43: 318–329.

Vasiliadis HM, Dionne PA, Preville M, Gentil L, Berbiche D, et al. (2013). The excess healthcare costs associated with depression and anxiety in elderly living in the community. Am J Geriatr Psychiatry 21: 536-548.

Vittengl JR, Clark LA, Dunn TW, Jarrett RB (2007). Reducing relapse and recurrence in unipolar depression: a comparative meta-analysis of cognitive-behavioral therapy's effects. J Consult Clin Psych 75(3): 475-488.

Von Neumann J, Morgenstern O (1953). Theory of games and economic behavior. Wiley, New York.

Vos T, Mathers CD (2000). The burden of mental disorders: a comparison of methods between the Australian burden of disease studies and the Global Burden of Disease study. Bull World Health Organ 78: 427–438.

Vos T, Haby MM, Barendregt JJ, Kruijshaar M, Corry J, et al. (2004). The burden of major depression avoidable by longer-term treatment strategies. Arch Gen Psychiatry 61(11): 1097-103.

Vos T, Haby MM, Magnus A, Mihalopoulos C, Andrews G, et al. (2005). Assessing cost-effectiveness in mental health: helping policy-makers prioritize and plan health services. Aust N Z J Psychiatry 39(8): 701-712.

Vos T, Carter R, Barendregt J, Mihalopoulos C, Veerman JL, et al. (2010). Assessing cost-effectiveness in prevention (ACE–Prevention): final report. University of Queensland, Brisbane and Deakin University, Melbourne.

Vos T, Flaxman AD, Naghavi M, Lozano R, Michaud C, et al. (2012). Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380(9859): 2163-2196.

Walters GD (2000). Behavioral self-control training for problem drinkers: a meta-analysis of randomized control studies. Behavior Therapy 31(1): 135-149.

Walters SJ, Brazier JE (2005). Comparison of the minimally important difference for two health state utility measures: EQ-5D and SF-6D. Qual Life Res 14: 1523–1532.

Ware JE, Sherbourne CD (1992). The MOS 36-item short-form health survey (SF-36): i. Conceptual framework and item selection. Med Care 3: 473–483.

Warmerdam L, Van Straten A, Twisk J, Riper H, Cuijpers P (2008). Internet-based treatment for adults with depressive symptoms: randomized controlled trial. J Med Internet Res 10: e44.

Weehuizen R (2008). Mental capital: the economic significance of mental health. University of Maastricht, Dissertation.

Wiles N, Thomas L, Abel A, Barnes M, Carroll F, et al. (2014). Clinical effectiveness and cost-effectiveness of cognitive behavioural therapy as an adjunct to pharmacotherapy for treatment-resistant depression in primary care: the CoBalT randomised controlled trial. Health Technol Assess 18(31): 1-167.

Wittchen HU (1994). Reliability and validity studies of the WHO. Composite International Diagnostic Interview (CIDI): a critical review. J Psychiatr Res 28: 57–84.

World Health Organization (1994). Lexicon of alcohol and drug terms. World Health Organization, Geneva.

World Health Organization (1997). Composite International Diagnostic Interview (CIDI), version 2.1, 12-months. World Health Organization, Geneva.

World Health Organization (2006). Dollars, DALYs and Decisions. World Health Organization, Geneva.

World Health Organization (2008). Mental Health Gap Action Programme: scaling up for mental, neurological and substance use disorders. World Health Organization, Geneva.

Willemse GR, Smit F, Cuijpers P, Tiemens BG (2004). Minimal contact psycho- therapy for sub-threshold depression in primary care: a randomised trial. Br J Psychiatry 185: 416-421.

Williams I, McIver S, Moore D, Bryan S (2008). The use of economic evaluations in NHS decision-making: a review and empirical investigation. Health Technol Assess 12(7): 1-175.

Wong A, Wouterse B, Slobbe LCJ, Boshuizen HC, Polder JJ (2012). Medical innovation and age-specific trends in health care utilization: Findings and implications. Soc Sci Med 74: 263-272.

Glossary

Glossary

This glossary of health economic terms presents a selection from the list of key terms presented on the websites of BMJ[2] and Centers for Disease Control and Prevention (CDC)[3].

Allocative efficiency
Occurs when, given the existing income distribution, resources cannot be reallocated so that they make one person better off (in terms of gaining greater satisfaction from the goods and services they consume) without making at least one other person worse off. This is also known as Pareto efficient.

Benefit
Anything that results that is of value.

Clinical effectiveness
The application of interventions which have been shown to be efficacious to appropriate patients in a timely fashion to improve patients' outcomes and value for the use of resources.

Cost
The economic definition of cost (also known as opportunity cost) is the value of opportunity forgone, strictly the best opportunity forgone, as a result of engaging resources in an activity. Note that there can be a cost without the exchange of money. Also the economists' notion of cost extends beyond the cost falling on the health service alone, e.g., includes costs falling on other services and on patients themselves.

Cost benefit analysis (CBA)
Expresses all gains and sacrifices in common units (usually money), allowing a judgement to be made of whether, or to what extent, an objective should be pursued.

---

[2] http://clinicalevidence.bmj.com/x/set/static/ebm/toolbox/678253.html. *"Taken from glossaries provided online by the All Wales Medicines Strategy Group and the US National Library of Medicine."*
[3] http://www.cdc.gov/dhdsp/programs/nhdsp_program/economic_evaluation/docs/Economic_Evaluation_Glossary.pdf

Cost effectiveness analysis (CEA)
An economic analysis in which all costs are related to a single, common effect. Results are usually stated as additional cost expended per additional health outcome achieved. Results can be categorized as average cost-effectiveness, marginal cost-effectiveness, and incremental cost-effectiveness.

Cost of illness study
Aims to identify and measure the total costs attributable to a particular disease. These are not a type of economic evaluation as they are not used to assess the costs and benefits of alternative interventions or programmes. They may provide useful information that can be used in the context of an economic evaluation of interventions related to the disease category, although care must be taken as not all costs included in a cost of illness study represent resource costs. Cost of illness studies may also be utilised in the estimation of the economic burden of disease.

Cost utility analysis (CUA)
A form of cost-effectiveness analysis where benefits are measured in terms of a utility measure such as the quality-adjusted life year (QALY).

Disability adjusted life years (DALYs)
A standard measure for comparing health outcomes for various health conditions; years of potential life lost due to premature mortality and the years of productive life lost due to disability.

Discount rate
The rate chosen to express the strength of preference over the timing of costs and benefits (see discounting and time preference).

Discounting
The most widely accepted method of incorporating time preference into the evaluation of a programme when the costs and benefits do not occur at the same point in time.

Economic evaluation (economic appraisal)
The comparison of alternative courses of action in terms of their costs and consequences, with a view to making a choice.

Effectiveness
The extent to which programmes achieve their objectives, in real-life settings.

Efficacy
The effect of an intervention under ideal conditions, with participants fully complying with the programme.

Efficiency
Maximising the benefit to any resource expenditure, or minimising the cost of any achieved benefit.

Equality
Equal shares of some good or service.

Equity
Fair distribution of resources or benefits among different individuals or groups.

Health economics
The study of how scarce resources are allocated among alternative uses for the care of sickness and the promotion, maintenance, and improvement of health, including the study of how health care and health-related services, their costs and benefits, and health itself are distributed among individuals and groups in society.

Health effects
These relate to specific outputs and outcomes resulting from a programme.

Incremental cost
The difference between the cost of a treatment and the cost of the comparison treatment.

Incremental cost-effectiveness ratio (ICER)
Obtained by dividing the difference between the costs of the two interventions by the difference in the outcomes, i.e., the extra cost per extra unit of effect.

Marginal analysis
The evaluation of the change in costs and benefits produced by a change in production or consumption of one unit. Less formally it is often used to refer to the change in costs and benefits produced by the particular change in scale of production or consumption which is under consideration.

Markov model
A particular type of decision analysis that allows for the transfer between different health states over a period of time.

Opportunity cost
The cost of a unit of a resource is the benefit that would be derived from using it in its best alternative use.

Outcome
The results and value of the intervention, e.g., intermediate measures such as number of quitters, or long-term outcomes such as life-years saved.

Output
The activities that result from the use of resources in the programme, e.g., number and type of materials given, number of client-professional contacts, and their type.

Perspective
The point of view from which an analysis is carried out. The social welfare perspective considers costs and benefits from the point of view of society.

Quality-adjusted life years (QALYs)
Calculated by adjusting the estimated number of life-years an individual is expected to gain from an intervention for the expected quality of life in those years. The quality of life score will range between 0 for death, to 1 for perfect health, with negative scores being allowed for states considered worse than death.

Resources
Things that contribute to the production of output. Money gives a command over resources but is not a resource per se.

Scarcity
There will never be enough resources to satisfy human wants completely.

Sensitivity analysis
A process through which the robustness of an economic model is assessed by examining the changes in results of the analysis when key variables are varied over a specified range.

Time preference
Individuals are not indifferent to the timing of costs and benefits, preferring benefits sooner and costs later.

Utility effects
In an attempt to generate measures which can be used to compare outcomes across all health care interventions, considerable effort has been invested in measures of health status and utility.

Willingness to pay
This technique asks people to state explicitly the maximum amount they would be willing to pay to receive a particular benefit. It is based on the premise that the maximum amount of money an individual is willing to pay for a commodity is an indicator of the value to them of that commodity.

Curriculum Vitae

CURRICULUM VITAE

Joran Lokkerbol obtained his Master degree in Econometrics in 2008 at the University of Amsterdam (specialization Mathematical Economics) and started his career as consultant at PricewaterhouseCoopers and the Institute for Business and Industrial Statistics. In 2011, Joran joined the Netherlands Institute of Mental Health and Addiction (Trimbos Institute), specializing in health economic evaluation and modelling. Later in 2011, Joran started working on his PhD at VU University, department of clinical psychology. In these years, he also founded kansvanslagen.nl, an online training in mathematics, Dutch and English, serving over 3,000 high school students in the last three years. In 2014, Joran became Head of the Economic Evaluation team at Trimbos Institute. In 2015 he received the *Excellence in Mental Health Policy and Economics Research Award*.

Joran's interests relate to economics, mathematics, and modelling, which he uses to combine the perspectives of health and economics by developing business cases and algorithms for health care system improvement. His health economic models are used by WHO, universities, governments, insurance companies and health care providers. Recently, Joran has been involved in the development of various multidisciplinary clinical guidelines.

Joran's current research is mostly related to the intersecting fields of economics and psychiatry. His previous work was published in the field of industrial statistics.

Publications

Peer-reviewed papers
Lokkerbol J, Adema D, Cuijpers P, Reynolds CF, Schulz R, Weehuizen R, Smit F (2014). Improving the cost-effectiveness of a healthcare system for depressive disorders by implementing telemedicine: a health economic modelling study. American Journal of Geriatric Psychiatry.

Lokkerbol J, Weehuizen R, Mavranezouli I, Mihalopoulos C, Smit F (2014). Mental health care system optimization from a health-economics perspective: where to sow and where to reap? Journal of Mental Health Policy and Economics 17: 51-60.

Kooistra LC, Wiersma JE, Ruwaard J, van Oppen P, Smit F, Lokkerbol J, Cuijpers P Riper H (2014). Blended vs. face-to-face cognitive behavioural treatment for major depression in specialized mental health care: study protocol of a randomized controlled cost-effectiveness trial. BMC Psychiatry 14: 290.

Lokkerbol J, Adema D, de Graaf R, ten Have M, Cuijpers P, Beekman A, Smit F (2013). Non-fatal burden of disease due to mental disorders in the Netherlands. Social Psychiatry and Psychiatric Epidemiology 48(10): 1591-1599.

Lokkerbol J, Does RJJM, de Mast J, Schoonhoven S (2012). Improving Processes in Financial Service Organizations: Where to Begin? International Journal of Quality and Reliability Management 29(9): 981-999.

De Mast J, Lokkerbol J (2012). An analysis of the Six Sigma DMAIC method from the perspective of problem solving. International Journal of Production Economics 139(2): 604-614.

Smit F, Lokkerbol J, Riper H, Majo MC, Boon B, Blankers M (2011). Modelling the cost-effectiveness of health care systems for alcohol use disorders: how implementation of eHealth interventions improves cost-effectiveness. Journal of Medical Internet Research 13(3): e56.

Non-peer reviewed papers / reports
Lokkerbol J, Elshaug A (2015). Disinvestment: a narrative review. Research Agenda for Health Economic Evaluation. WHO, Brussels.

Lokkerbol J, Lokman S, Boon B, Smit F (2014). Kosten en effecten van preventie van middelengebruik bij jongeren. Trimbos-instituut, Utrecht.

Lokkerbol J, Molenaar M, Does RJMM (2012). Quality Quandaries: An efficient public sector. Quality Engineering 24: 431–435.

Lokkerbol J, van Splunteren P, Lokman S, Smit F (2012). Versterken van mentaal kapitaal in het onderwijs. Trimbos-instituut, Utrecht.

Lokkerbol J, van Splunteren P, Conijn B, Smit F (2012). Versterken van mentaal kapitaal in de gemeente. Trimbos-instituut, Utrecht.

Lokkerbol J, van Splunteren P, Smit F (2012). Versterken van mentaal kapitaal in de werksetting. Trimbos-instituut, Utrecht.

Lokkerbol J, van Splunteren P, Smit F (2012). Versterken van mentaal kapitaal in de zorg. Trimbos-instituut, Utrecht.

Lokkerbol J, Schotman MAM, Does RJMM (2011). Quality Quandaries: Personal injuries: a case study. Quality Engineering 24(1): 102-106.

Lokkerbol J, Smit F (2011). Eigen bijdrage in GGZ: scenario's van effecten op zorguitgaven, zorggebruik en arbeidsproductiviteit. Trimbos-instituut, Utrecht.

Lokkerbol J, Verhaak P, Smit F (2011). Op zoek naar een kosteneffectieve basis GGZ: scenario's voor zorginnovatie. Trimbos-instituut, Utrecht.

In preparation
Kleiboer A, Lokkerbol J, Donker T, van Straten A, Mihalopoulos C, Cuijpers P, Smit F: The role of support in web-based treatment for depression and anxiety in adults: cost-utility analysis alongside a randomized trial.

Lokkerbol J, Weehuizen R, Cuijpers P, Mihalopoulos C, Smit F. Improving the cost-effectiveness of health care by simultaneous investing and disinvesting in interventions.

Biesheuvel-Leliefeld KEM, Lokkerbol J, Smit F. Cost-effectiveness of preventing recurrences in depressive disorder: a population-level health economic modelling study.

Lokkerbol J. Estimating incidence using prevalence estimates during a variable time frame and episode duration.

Book
De Mast J, Does R, de Koning H, Lokkerbol J (2012). Lean Six Sigma for services and healthcare. Beaumont, Alphen aan den Rijn.

Acknowledgement

Acknowledgement

This dissertation would never have been written without the help of the many people who were involved, either directly or indirectly. Above all, I owe much gratitude to my first thesis supervisor, prof.dr. Filip Smit (professor of public mental health, VU University). Filip, working with you simply is a pleasure. Your patience, kindness, expertise and positive attitude are a continuous source of inspiration. There is too much I could say here, but let me specifically thank you for the pivotal role you played, both in the creation of this thesis as well as the road that led to starting it.

Many thanks go out to my second thesis supervisor, prof.dr. Pim Cuijpers (professor of clinical psychology, VU University), for letting me write this thesis under the flag of his department, for allowing me to dedicate substantial time to writing this thesis, and above all, for a never-ending good spirit and willingness to contribute to whatever is needed.

Dr. Cathy Mihalopoulos (associate professor, Deakin University) and dr. Rifka Weehuizen (managing director, University of Strasbourg Institute for Advanced Study); your work in the area of mental health, your vast knowledge in the field of economic theory and your creative minds served as important sources of inspiration for this thesis. I am honoured to have worked on this with both of you.

Special thanks to the members of the reading committee: Annemieke van Straten (professor of clinical psychology, VU University), Patrick Jeurissen (chief strategy group, Ministry of Health; chief program fiscal sustainability, Radboud University Medical Centre), Richard Janssen (professor of economics and organization of health care, Tilburg University), Robert Schoevers (professor of psychiatry and head of the psychiatry department, University Medical Centre Groningen), Silvia Evers (professor in public health technology assessment, Maastricht University), and Veerle Coupé (assistant professor, VU University) for taking the time to read the manuscript carefully and providing me with valuable comments.

By viewing the healthcare system predominantly from a macro-perspective, I merely connected the dots between large bodies of existing knowledge. This thesis could not have existed without the great efforts that many, many people before me undertook to further science in the fields of epidemiology, clinical effectiveness and health economics. I was lucky to find myself working with a number of these people. For valuable contributions made as co-author, I would like to give many thanks to Aartjan Beekman, Annet Kleiboer, Brigitte Boon, Charles Reynolds, Cristina Majo, Dirk

216

Thank you Ronald Does and Jeroen de Mast, for introducing me to the joy of publishing.

Writing this thesis next to my regular job at Trimbos over a period of three years, as well as setting up a company and playing in a band, would never have been possible without the support of those around me. I feel very lucky to be surrounded by many kind-hearted, talented and friendly people.

Roderick, thank you for the friendship, for the camping trips, for being my personal trainer. Arshia, mate, your sincere kindness is a blessing to everyone around you; I can only feel extremely grateful to count you as my friend.

Thank you Bart, Emiel, Dima for sharing music and friendship, for sharing unique adventures in even more unique settings and for putting up with my priorities more often than not drifting off to work. Dirk, perhaps unknowingly, you played a very big role in where I stand today. Besides that, your positive attitude, sense of humour and taste of music make you one of the more pleasant people to spend time with. Let's do that more often. René; it goes without saying that I'm very happy to count you as one of my friends. There are many reasons for that, the shear abundance of shared interests being just one of them. Elwin, Jochem, Pieter, Ramon; thanks for not taking this all too seriously. Robert, Pieter, Kasper, thanks for the fun and the experience. Rogier, thank you for providing me with a roof throughout quite a significant part of my life, as well as all the other things you did. Anna, Bas, Bastiaan, Benjamin, Jasper, Maarten, Martijn, Mieke, Niels, Sander, Tashi, thank you for your friendship.

Needless to say, my family played an important role by supporting me and taking care of me throughout. Mam, Pim, Camiel, Christa, Enno, Bianca, Nadine; each of you is more inspiring to me than you probably realize. Dad, this work could never have existed if it were not for your lessons of perseverance and discipline. Thijs, Lise, Jasmijn: with about 6 years of life experience, in total, you already excel in putting projects such as this thesis into a healthy perspective. Oma, mijn studie is nu echt af, bedankt dat uw deur altijd open staat! Maartje, Peter, Ier, Vin, Roos, Lucas and oma; thank you for letting me be a part of the family, including all the delicious food, wonderful trips and good conversations.

And most of all, Eva, for growing with me through life, for building a home and growing a family together, for supporting me throughout with endless amounts of patience, and for never ceasing to remind me of the important things in life.

Meeting the needs of the many patients suffering from mental disorders requires a qualitatively good health care system. However, financial budgets for mental health care are under pressure.

In this context, the following questions need to be addressed urgently:

- How can cost-effective interventions be identified?
- Will the health care system become more cost-effective when e-health is being introduced on a large scale? How will this affect population health and budgets?
- And in an award-winning paper: what interventions need to be scaled up to increase the overall cost-effectiveness of a health care system? Can we strike an optimal balance between investing and disinvesting in interventions?

This thesis presents both existing and new economic evaluation methods to answer these and related questions.

Joran Lokkerbol is head of the Economic Evaluation team at the Netherlands Institute of Mental Health and Addiction (Trimbos Institute). He works to optimize and rationalize decision-making in the mental health care field.