EXPERT REVIEW OF PHARMACOECONOMICS & OUTCOMES RESEARCH
2021, VOL. 21, NO. 5, 1031–1042
https://doi.org/10.1080/14737167.2021.1844566

Taylor & Francis
Taylor & Francis Group

ORIGINAL RESEARCH

Design of a health-economic Markov model to assess cost-effectiveness and budget impact of the prevention and treatment of depressive disorder

Joran Lokkerbol [a], Ben Wijnen [a,b], Henricus G. Ruhe[c,d,e], Jan Spijker[f,g], Arshia Morad[h], Robert Schoevers[c,i], Marrit K. de Boer[c], Pim Cuijpers [j,k] and Filip Smit[a,j,k]

[a]Centre for Economic Evaluation and Machine Learning, Department of Public Mental Health, Trimbos Institute (Netherlands Institute of Mental Health and Addiction), Utrecht, The Netherlands; [b]Department of Clinical Epidemiology and Medical Technology Assessment, Maastricht University Medical Centre+, Maastricht, The Netherlands; [c]Department of Psychiatry, University of Groningen, University Medical Center Groningen, Groningen, The Netherlands; [d]Radboudumc, Department of Psychiatry, Radboud University of Nijmegen, Nijmegen, The Netherlands; [e]Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, The Netherlands; [f]Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands; [g]Depression Expertise Centre, Pro Persona Mental Health Care, Nijmegen, The Netherlands; [h]School of Psychology, The University of Sydney, New South Wales, Australia; [i]Research School of Behavioural and Cognitive Neurosciences (BCN), Interdisciplinary Center for Psychopathology and Emotion Regulation (ICPE), Groningen, The Netherlands; [j]Department of Clinical, Neuro and Developmental Psychology, Amsterdam Public Health Research Institute, Vrije Universiteit Amsterdam, The Netherlands; [k]Department of Epidemiology and Biostatistics, Amsterdam Public Health Research Institute, Academic Medical Centers Amsterdam, Location VUmc, Amsterdam, The Netherlands

ABSTRACT

Background/objective: To describe the design of 'DepMod,' a health-economic Markov model for assessing cost-effectiveness and budget impact of user-defined preventive interventions and treatments in depressive disorders.

Methods: DepMod has an epidemiological layer describing how a cohort of people can transition between health states (sub-threshold depression, first episode of mild, moderate or severe depression (partial) remission, recurrence, death). Superimposed on the epidemiological layer, DepMod has an intervention layer consisting of a reference scenario and alternative scenario comparing the effectiveness and cost-effectiveness of a user-defined package of preventive interventions and psychological and pharmacological treatments of depression. Results are presented in terms of quality-adjusted life years (QALYs) gained and healthcare expenditure. Costs and effects can be modeled over 5 years and are subjected to probabilistic sensitivity analysis.

Results: DepMod was used to assess the cost-effectiveness of scaling up preventive interventions for treating people with subclinical depression, which showed that there is an 82% probability that scaling up prevention is cost-effective given a willingness-to-pay threshold of €20,000 per QALY.

Conclusion: DepMod is a Markov model that assesses the cost-utility and budget impact of different healthcare packages aimed at preventing and treating depression and is freely available for academic purposes upon request at the authors.

ARTICLE HISTORY
Received 2 September 2020
Accepted 28 October 2020

KEYWORDS
Cost-effectiveness; budget impact; major depressive disorder; health-economic modeling; depression

1. Introduction

Depressive disorder has consistently been highlighted as a leading cause of disease burden, particularly in terms of years lived with disability (YLD) [1–6]. Depression is associated with substantial economic costs from a patient perspective, a healthcare perspective and a societal perspective [7,8]. For example, the total annual cost of depression in Europe was estimated at €118 billion in 2004 [8].

Extensive research has been done on the effectiveness of both the treatment of depression (e.g. [9–13]), and the prevention of depression [14,15]. Optimizing a healthcare system under restricted budgets requires that the cost-effectiveness of a package of interventions can be assessed along with associated impacts of implementing an intervention on the healthcare budget. Such information helps to inform policymakers about the implications of reforming the healthcare system. One way of addressing these issues is to develop a health-economic model which synthesizes all available evidence, while also considering preferences of patients and healthcare professionals.

The aim of this paper is to present a health-economic simulation model for depression, DepMod, and to describe DepMod's design features and how it can be put to use to inform policymakers about the health and cost implications of changing the healthcare system for depressive disorders. DepMod builds on previous health-economic models, such as the ACE (Assessing Cost Effectiveness) prevention models [16], and CHOICE (CHOosing Interventions that are Cost-Effective) models [17]. To promote transparency regarding the model structure (its inputs, assumptions, strengths and limitations), this paper presents the building blocks of the model to enable future users of the model to use it for

CONTACT Ben Wijnen bwijnen@trimbos.nl Centre of Economic Evaluation (Trimbos Institute), Netherlands Institute of Mental Health and Addiction, Utrecht, The Netherlands

Article highlights

- DepMod is a methodologically sound model that can be used to examine the cost-effectiveness and budget impact of a user-defined package of preventive interventions and psychological and pharmacological treatments of depression
- DepMod can be used for various interventions (e.g., medication, cognitive behavior therapy, et cetera) in a relatively easy and accessible way.
- The transition rates between health states were derived from NEMESIS-2, a large psychiatric cohort study of adults (18 – 65 years) in the Netherlands, but DepMod permits user-defined adaption of its epidemiology for use in other geographies or age groups.
- For illustrative purposes, DepMod was used to assess the cost-effectiveness of scaling up preventive interventions for treating people with subclinical depression in the Netherlands, which showed that there is an 82% probability that scaling up prevention is cost-effective given a willingness-to-pay threshold of €20,000 per QALY.
- DepMod is freely available for academic purposes upon request by the authors.

answering their own research questions and adapt the model to their own (geographical) context in terms of epidemiology and treatment mix for the prevention and treatment of major depressive disorder.

2. Methods

2.1. Problem definition and target population

The aim of the model is to compare the cost-effectiveness and budget impact of competing intervention packages targeting adults with, or at risk of developing, a depressive disorder in the Netherlands, such that the model contributes to optimizing health outcomes under budget constraints.

2.2. Model development

DepMod was developed as a decision-analytic and health-economic Markov model built-in Microsoft Excel. Such a model evaluates costs and health outcomes by simulating a cohort of hypothetical patients that transition across a series of health and disease states (healthy, at risk, depressed, relapsing, chronic (partially) recovered, death) until the end of the model's time horizon. The outcomes of the model (costs and quality-adjusted life years gained) are then synthesized into incremental cost-effectiveness ratios [18].

The development of the health-economic model DepMod was in line with the recommendations from the International Society for Pharmacoeconomics and Outcomes Research [19], starting off with scoping the types of research questions that DepMod should be able to answer, conceptualizing the model, gathering evidence from relevant sources, and testing the model with the model users and stakeholders. Figure 1 depicts the schematic overview of the development of the model.

Expert input was used throughout the process of developing the health-economic model. In drafting the starting point of the conceptual model, researchers in the field of depression were consulted.

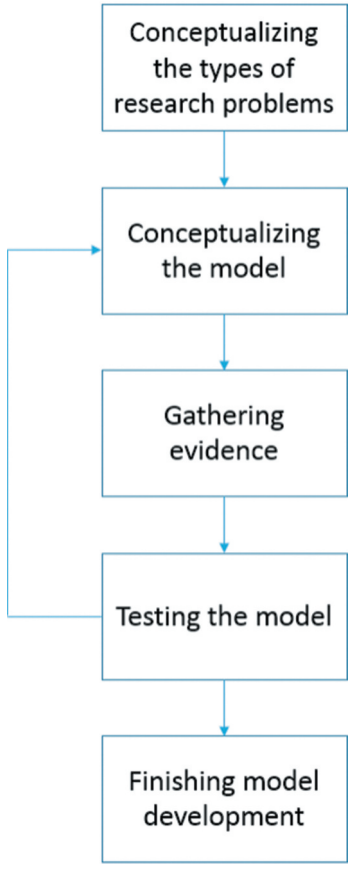

Figure 1. Schematic overview of the model development process.

The model was then refined in an iterative process where the model was updated multiple times based on feedback by healthcare users, healthcare professionals, and researchers involved in the development of a Standard of Care for treating depressive disorders in the Netherlands [20].

During the iterative updating process, DepMod's scope was defined in terms of the following PICOT:

- *Population*: Dutch adults (aged 18–65 years) with subthreshold, mild, moderate, and severe major depression as defined by DSM-IV with the possibility of recurrences, which is in line with the major source of available epidemiological evidence (i.e. The Netherlands Mental Health Survey and Incidence Study-1 (NEMESIS-1), and NEMESIS-2 [21–23]).
- *Intervention*: User-defined intervention(s) which will be compared to a reference scenario. The intervention in DepMod can either be a single intervention or a mix of interventions (e.g., scaling up existing evidence-based interventions or adding new preventive interventions).
- *Comparator*: User-defined reference scenario, typically the treatments provided under care as usual, to which the intervention will be compared.
- *Outcome*: Additional healthcare expenditure (in €) relative to quality-adjusted life years (QALYs) gained.
- *Time horizon*: Costs and QALYs can be modeled out over 5 years (as there is a paucity of evidence in the literature on longer-term health effects induced by treatments).

In addition, the following a priori design criteria for the model were adopted:

- Taking a Markov model approach similar to ACE [16] and CHOICE [17]. A Markov model was preferred over a micro-simulation (or discrete event) approach, as our aim was to make a population-level model, which makes individual traits affecting the course of illness less relevant.
- Taking a healthcare system approach, looking at packages of interventions (instead of individual interventions) to be compared to a 'reference scenario' (e.g., depression care concordant with the latest clinical guideline) with an 'alternative scenario' (e.g., an alternative user-defined package of interventions for depression).
- Using an incidence-based model (e.g., of the depressive disorder in the Dutch adult population), to be able to examine the impact of preventive interventions.

2.3. A layered model structure

DepMod is a layered model. In the first layer, the epidemiology of depression in the Dutch population is simulated (incidence, disease duration, remission, recurrence, chronicity, and depression-related excess mortality). A second layer is superimposed on the epidemiological substratum. Here a set of interventions can be modeled along with their coverage and compliance rates, their effectiveness and the costs of offering the interventions. Both layers are described in more detail below.

2.3.1. Epidemiological layer

To apply DepMod to the Dutch situation, NEMESIS-1 and NEMESIS-2 [21–23], both Dutch population-based cohort studies on the incidence and prevalence of mental disorders, served as the main sources to build DepMod's epidemiological substratum. NEMESIS-2 was used to obtain the most recent available data on the incidence and prevalence of depression. Furthermore, experts provided literature on chronicity [24] (partial) recovery/duration [25], and recurrence of depression [26]. These were modeled as transition rates between the health and disease states In DepMod. The model takes a population-incidence approach, where available evidence on the incidence of depression was used as direct input to the model, and transition parameters such as remission and recurrence rates were calibrated such that resulting prevalence-based outcomes matched the prevalence-based outcomes reported in NEMESIS-2.

2.3.2. Intervention layer

Superimposed on the epidemiological layer, DepMod has an intervention layer consisting of a reference scenario and an alternative scenario to compare the effectiveness and cost-effectiveness of a user-defined package of preventive interventions and psychological and pharmacological treatments of depression. Each intervention is described by four parameters: effectiveness (standardized mean difference in case of treatment, and risk difference or relative risk in case of prevention), per-patient intervention costs (in €), coverage (the fraction of the target population receiving the intervention) and adherence (the fraction of users compliant with that intervention).

End-users can modify the included interventions in DepMod according to their own preference. The effectiveness parameters for interventions are best extracted from the meta-analytical trial literature, while coverage and adherence rates can be elicited from focus groups of clinicians and patients, respectively.

2.3.3. Conceptual model

Figure 2 depicts the conceptual model formed by depression experts that was used as a starting point in developing the model. The conceptual model begins with the population at-risk of developing a depression (e.g., the estimated 456,600 people with subthreshold depression (warranting indicated prevention) in the Dutch adult population), which can develop into a mild, moderate or severe major depressive disorder. Depression can then either remit (fully or partially) or not, and -in case of (partial) remission-transit into a recurrent major depressive disorder or remain a single episode. People developing a second episode of depression reenter the loop, where they can again remit and develop a next episode of depression.

2.4. Perspective, outcomes and time horizon

DepMod takes a healthcare perspective, considering the costs and effects of interventions for preventing first episodes of depression, treatment of acute depressive episodes, and prevention of recurrent episodes of depression. Health outcomes are expressed as QALYs. QALYs are based on the health state valuations for mild, moderate, and severe depressive disorder of 0.86, 0.65, and 0.24, respectively, as reported by Stouthard et al. [27], whereas changes in health-related quality of life are based on the standardized effect sizes of the interventions that are converted to utilities using Sanderson et al.'s (2004) conversion factor [28]. Total QALYs are estimated by multiplying the time spend in a specific health state by the valuation (utility) of that health state [18]. The model uses a time horizon of 5 years to capture the longer-term health effects of treatment and prevention of depression without extrapolating too far from the available evidence-base. Costs and QALYs occurring after 1 year are discounted at 4% and 1.5%, respectively, to account for differential timing, in accordance with the Dutch guideline for health economic evaluation [29].

In DepMod, the incremental cost-effectiveness ratio for the alternative scenario as compared to the reference scenario, is calculated by dividing the difference in total healthcare costs (between the alternative and reference scenario) by the difference in effects between the scenarios. Probabilistic sensitivity analyses are conducted by calculating the costs and effects in the reference scenario and alternative scenario by a user-defined number of times (e.g., 1,000 times) using Monte Carlo simulation, where each time a parameter value for each of the parameters is drawn at random from underlying cost, effect, and transition distributions. We assume costs to be Gamma distributed, standardized effect sizes to be

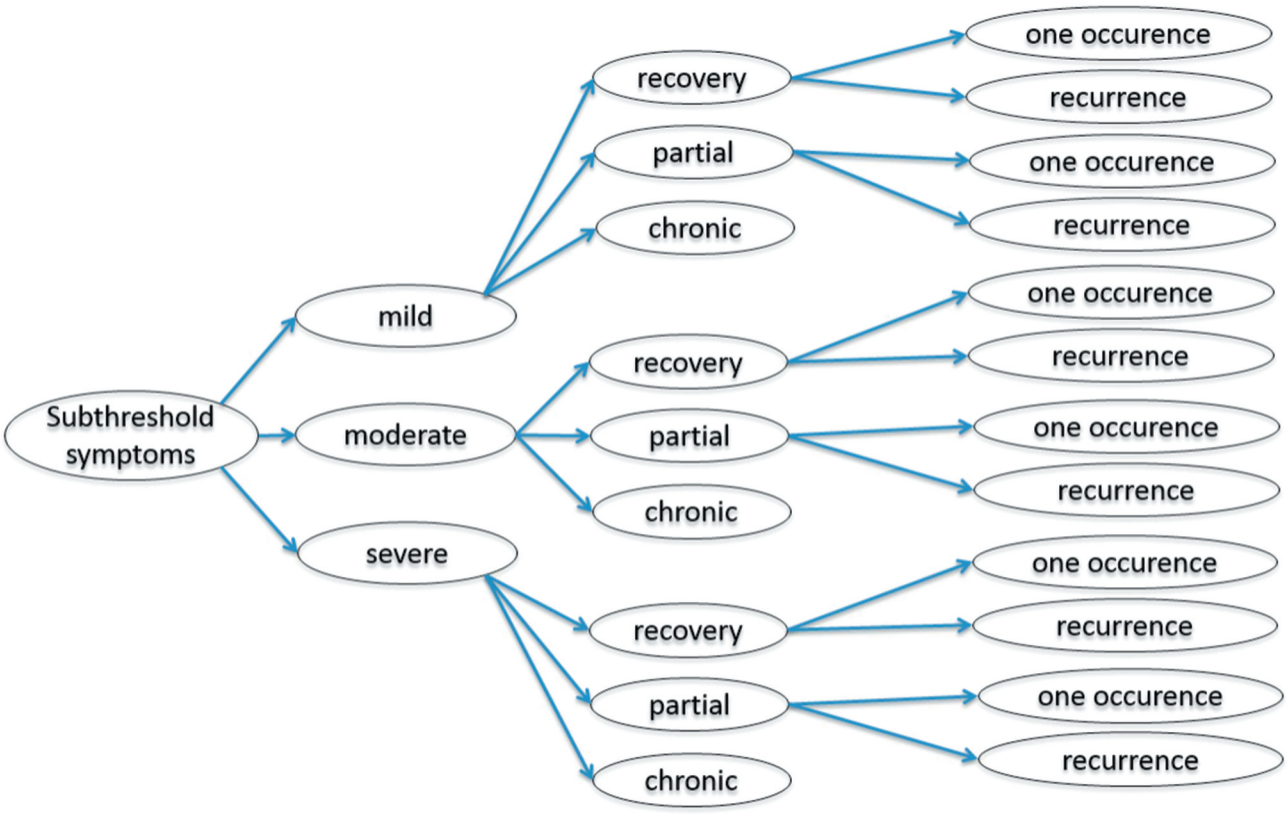

Figure 2. Conceptual model of the course of depression serving as a starting point in the process of model development.

normally distributed, and transition probabilities to follow a beta-distribution, in line with recommendations [18]. Results of probabilistic sensitivity analyses are depicted in a cost-effectiveness plane and a cost-effectiveness acceptability curve, presenting the probability that the alternative healthcare scenario is cost-effective compared to the reference scenario, given varying levels of willingness-to-pay for a QALY gained [30].

2.5. Epidemiology and the Markov model

Available evidence on epidemiology determined to which extent the level of detail as depicted in the conceptual model in Figure 2 could be modeled. As available evidence reported mostly in time intervals of 1 year, it was chosen to use a cycle length in our Markov-model of 1 year as well. A half-cycle correction was applied to account for the fact that transitions between states are expected to occur on average in the middle of each cycle [31]. It was chosen to only model the states 'subthreshold depression' (without a history of depression), 'depressive episode' (first or recurrent episode), 'no depression,' a 'chronic state' (i.e., more than 2 years), 'cured' (absence of (subthreshold) depression, and 'death.' In line with Solomon et al. (2000), up to five depressive episodes were considered, which resulted in the Markov-model as depicted in Figure 3 [32]. Technically, to allow for increasing recurrence rates after each previous episode (i.e., to build memory into the Markov-model), we constructed a transition matrix containing tunnel states [33] for the number of depressive episodes (at most five, i.e., one per year), representing the transition rates after no previous episode, one previous episode, etc.

Based on NEMESIS-2, the first (one-year) incidence rate was set at 1.28% [23]. Transition parameters that could not be taken from literature were calibrated (see Table 1) such that the resulting epidemiology matched the internally consistent epidemiological structure of depression as derived from the NEMESIS-studies (see Figure 4). Recurrence rates were based on Solomon et al. (2000), who reported that the risk of recurrence increases with on average 16% after each previous episode of major depression. In our model, recurrence rates ranged from a 21% risk of recurrence within 12 months given one previous episode, to an 88% risk of recurrence within 5 years given four previous episodes [32]. Solomon et al. (2000) report 13 recurrence rates (for differing number of previous episodes and number of years until recurrence), along with their confidence intervals. DepMod defines 25 distinct recurrence rates (one to five recurrences that can each occur 1 to 5 years after the previous episode), which was simplified by estimating the average relative increase in recurrence rates after the first up to the fifth episode (so five increased recurrence rates) and five probabilities regarding the number of years after which the recurrence occurs. To allow for increasing recurrence rates, each of the depressive episodes (first to the fifth) was modeled as separate Markov tunnel states.

In line with NEMESIS-2, reporting a prevalence of major depression of 5.2% [34], the epidemiology as modeled by DepMod resulted in a 12-month prevalence rate of 5.18%. Chisholm et al. (2004) reported the proportion of people with a mild, moderate, and severe depression as 30%, 47%,

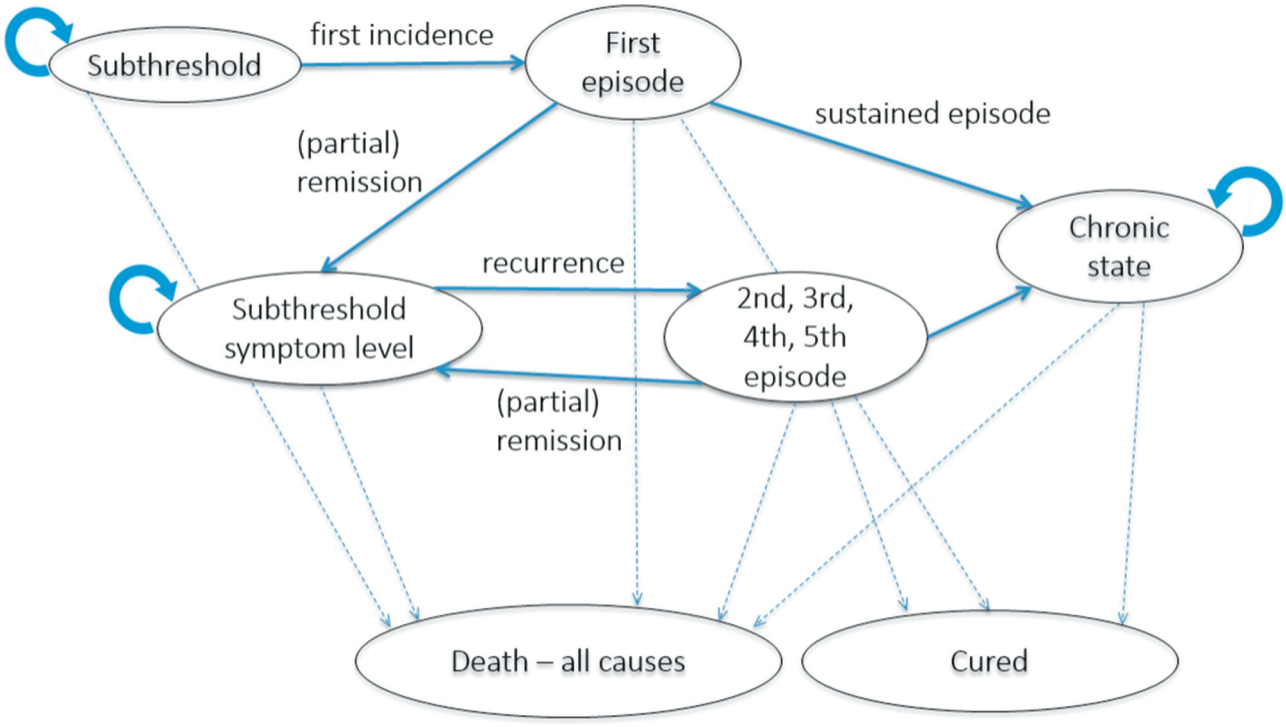

Figure 3. Markov-model.

Arrows representing all possible transitions from one state to another. In the model, each arrow is represented by a transition probability.

and 23%, respectively [35]. The epidemiology in DepMod was calibrated to match this distribution (see Table 1).

Background mortality was based on (weighted) average mortality rate of adults between 18 and 65 in the Netherlands [36]. Depression-related excess mortality (e.g., due to unhealthy behavior or suicide) was calibrated to be in line with NEMESIS-2, resulting in a relative risk of 1.65.

An economic evaluation compares the costs and effects of two alternatives. In DepMod, the reference scenario can be compared to an alternative scenario, simply by copying in this reference scenario and adjusting some of its parameters. For example, increasing the coverage rate of prevention implies that more prevention is offered in the alternative scenario, which may then cascade into a range of health and economic impacts (see, e.g. [37]).

2.6. Effectiveness of interventions

The effectiveness of interventions impacts positively on the epidemiology in three different ways:

- through a reduction in the transition from the at-risk status to a first episode of depression (primary prevention) and a reduction in the transition to a recurrent depression (prevention of recurrence). The effectiveness of prevention is expressed as either a risk difference (primary prevention) or a relative risk (prevention of recurrence).
- through symptom improvement in people currently suffering from mild, moderate or severe depression. Symptom improvement follows from standardized effect sizes (Cohen's *d;* i.e. standardized mean difference), expressing improvement on a depression-related symptom scale in standard units, and should be based on meta-analyses where possible. The average symptom improvement can be downward adjusted for less than optimal coverage and impaired adherence (by multiplication with the coverage and adherence rates). The resulting adjusted effect sizes are then converted into QALYs using Sanderon et al.'s (2004) conversion factor [28] and assuming that treatment effects would last as long as (but not longer than) the average episode duration of depression.
- through a prophylactic effect, where receiving psychological intervention for the treatment of a depressive episode may increase the likelihood of staying in remission after successful treatment (rather than recurrence at end of treatment), because psychological intervention (unlike pharmaceutical intervention) is suggested to train patients to better self-manage emerging depressive symptoms should these re-occur in the future.

The total QALY gain is then determined by difference in total QALYs between the reference and alternative scenarios.

3. Results of model testing

To apply DepMod to a real-world example, a list of evidence-based interventions needs to be constructed. In our application, we constructed a reference scenario by using a list of

Table 1. Parameters in DepMod (point estimates, range, justification).

Input parameters	Point estimate	Sensitivity	Source
Incidence	1.28%	NA	De Graaf et al., 2013[23]
Disability weights depression			
Mild (dw)	0.14	0.086 - 0.194	Stouthard et al., 1997[27]
Moderate (dw)	0.35	0.272 – 0.425	Stouthard et al., 1997[27]
Severe (dw)	0.76	0.556 – 0.971	Stouthard et al., 1997[27]
Depression duration			
Mild	3 months	NA	based on Spijker et al., 2002[24]
Moderate	6 months	NA	based on Spijker et al., 2002[24]
Severe	9 months	NA	based on Spijker et al., 2002[24]
Background mortality	0.001984	NA	Based on life tables [36]
Calibration parameters	**Point estimate**	**Sensitivity**	**To match outcome**
severity distribution at incidence			
Mild	30.0%	NA	distribution prevalence mild, moderate, severe depression as reported in Chisholm et al., 2004 [35]
Moderate	47.0%	NA	
Severe	23.0%	NA	
Recurrence rates			
after 1st episode	59.8%	NA	recurrence rates as reported in Solomon et al., 2000[32]
after 2nd episode	77.2%	NA	
after 3rd episode	84.2%	NA	
after 4th episode	91.1%	NA	
after 5th episode	92.9%	NA	
Occurring after			
one year	40%	NA	recurrence rates as reported in Solomon et al., 2000[32]
two years	30%	NA	
three years	20%	NA	
four years	7%	NA	
five years	3%	NA	
Excess mortality	1.65	NA	De Graaf et al., 2012[34]
Outcomes	**Model outcome**	**Reference value**	**Source**
Prevalence of depression	5.24%	5.21%	De Graaf et al., 2012[34]
Severity distribution			
Mild	30.4%	30%	Chisholm et al., 2004[35]
Moderate	47.0%	47%	Chisholm et al., 2004[35]
Severe	22.7%	23%	Chisholm et al., 2004[35]
Recurrence rates	ranging from 24% (risk of recurrence after 1 year, given 1 previous episode) to 84% (risk of recurrence after 5 years given 4 previous episodes)	ranging from 25% (20%-32%) (risk of recurrence after 1 year, given 1 previous episode) to 79% (35%-95%) (risk of recurrence after 5 years given 3 previous episodes)	Solomon et al., 2000[32]

evidence-based interventions, using the Dutch multidisciplinary guideline for depression (see Table 2) [20]. This healthcare system served as a reference scenario against which alternative intervention mixes (i.e., alternative scenarios) could be compared in terms of cost-effectiveness and budget impact. The effectiveness of included interventions was based on meta-analyses (see Table 2).

With the interventions in the reference scenario defined, it is possible to conduct what-if analyses, where the user can assess the cost-effectiveness of hypothetical alternative healthcare systems. For example, currently, first-incidence prevention and prevention of depression is not systematically offered in the Netherlands although it has been recommended in clinical guidelines for some time. For illustrative purposes, we've assessed the cost-effectiveness of scaling up preventive interventions for people with subclinical depression (i.e., online CBT) and people at risk of developing a recurrent depression with a coverage of 15%. Hence, we simply copied the reference scenario as outlined in Table 2 into the alternative scenario and increased the coverage rates of the preventive interventions from 0% to 15%.

In this example, intervention costs were estimated by multiplying the guideline concordant resource use associated with each intervention by their standard (unit) cost price, as listed in the Dutch guideline for health-economic evaluation [29]. Costs were expressed in 2019 Euros. The resulting costs per intervention are shown in Table 3.

Increasing the coverage rate of prevention from 0% to 15% and subsequently simulating the costs and effects of both healthcare systems, results in the cost-effectiveness plane and cost-effectiveness acceptability curve depicted in Figure 5. The output tells us that at the commonly accepted willingness-to-pay threshold of €20,000 per QALY, there is an 82% probability that scaling up prevention is cost-effective.

Furthermore, the expected 18,400 QALYs gained by offering the additional prevention over the considered time horizon of 5 years require an additional budget of €191.1 million euros (net present value). Tabulated results can be found in Appendix 1.

4. Discussion

We built a Markov model to evaluate the cost-effectiveness and budget impact of a healthcare system of psychological and pharmacological interventions for people with depressive disorders, as well as people at risk of developing a first or

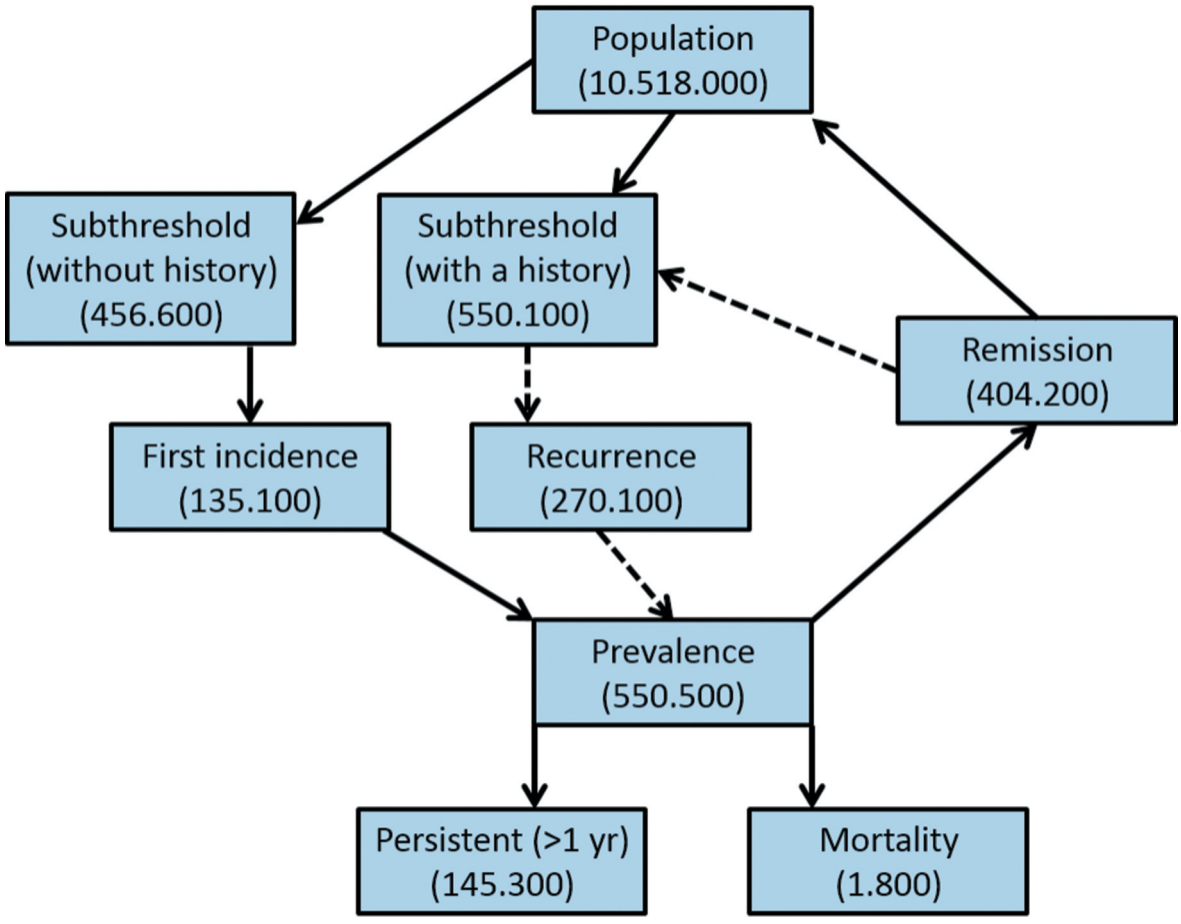

Figure 4. Internally consistent epidemiological structure of depression based on NEMESIS-studies (yearly number of people in each health state in parentheses).

Table 3. Intervention costs expressed in 2019 Euros.

Intervention by depression severity level		
Prevention of first incidence (targeting subclinical depression)	*€**	*Resource use*
E-health intervention (unsupported)	*160*	*1x GP; intake psychologist; 4–9 sessions unsupported online self-help; incl. hosting costs*
Treatment of mild depression	*€*	*Resource use*
E-health intervention (supported)	*265*	*1x GP; intake psychologist; 4–9 sessions online self-help; 4–5 telephone support prevention worker; hosting cost of the e-health intervention*
Individual psychotherapy	*962*	*1x GP; 8 sessions with a psychologist; coordination with GP*
Treatment of moderate depression	*€*	*Resource use*
E-health intervention (supported)	*265*	*1x GP; intake psychologist; 4–9 sessions online self-help; 4–6 telephone support prevention worker; hosting*
Individual psychotherapy	*962*	*1x GP; 8 sessions with a psychotherapist; coordination with GP*
Treatment of severe depression	*€*	*Resource use*
Individual psychotherapy, outpatient care	*1,984*	*1x GP; 8–24 sessions by a psychotherapist*
Antidepressants	*607*	*12 months of medication; contacts with GP or psychiatrist*
Antidepressants with additional psychological support	*679*	*12 months of medication; contacts with GP or psychiatrist; 3–6 visits GP support*
Combination therapy (medication plus psychotherapy)	*2,520*	*12 months of medication; contacts with GP or psychiatrist; 8–24 sessions with a psychotherapist*
Prevention of recurrent depression	*€*	*Resource use*
Clinical management with maintenance medication	*1,002*	*1x psychiatrist; 8–14 visits GP; 12 months of medication*
Mindfulness-based CT	*361*	*1x GP; 8 group sessions (11 participants on average)*
Cognitive (behavior) therapy	*429*	*1x GP; 8 group sessions (8 participants on average)*

* Intervention costs were estimated by multiplying the guideline concordant resource use associated with each intervention by their standard (unit) cost price, as listed in the Dutch guideline for health-economic evaluation.

recurrent episode of depression. Taking the healthcare system perspective, the model combines available evidence on the epidemiology of depressive disorder, the effectiveness of interventions, and the costs of the interventions while considering the preferences of healthcare professionals and healthcare users for offering and receiving interventions. The

Table 2. Illustration of selected evidence-based interventions by depression severity level: target group reached by the intervention expressed as Coverage (%) and Compliance with therapy (%). Effect expressed as risk difference (RD) or Relative Risk (RR) when impacting on transitions or as standardized effect size (Cohen's *d*) when impacting on symptom severity, all representing average values.

Intervention by depression severity level	Coverage rate	Compliance rate	Effect
Prevention of first incidence (targeting subclinical depression)	*%*	*%*	*RD*
E-health intervention (unsupported) [1]	*0*	*80*	*0.077*
Treatment of mild depression	*%*	*%*	*d*
E-health intervention (supported) [2]	*2*	*43*	*0.33*
Individual psychological intervention, primary care, 8 sessions [3]	*17*	*56*	*0.51*
Treatment of moderate depression	*%*	*%*	*d*
E-health intervention (supported) [2]	*2*	*43*	*0.33*
Individual psychological intervention, primary care, 8 sessions [3]	*16*	*56*	*0.51*
	%	*%*	*d*
Individual psychological outpatient care, 8–24 sessions [3]	*18*	*68*	*0.51*
Anti-depressants, 3–6 months via GP [4]	*20*	*44*	*0.30*
Anti-depressants, 3–6 months, with psychological support [3]	*20*	*56*	*0.51*
Prevention of recurrent depression	*%*	*%*	*RR*
Clinical management with maintenance medication, 12 months [5]	*0*	*42*	*0.75*
Mindfulness-based CT [6]	*0*	*68*	*0.66*
Cognitive (behavior) therapy [6]	*0*	*56*	*0.68*

[1]Taken from van Zoonen et al. 2014 [15]; [2] Taken from Karyotaki et al. 2017 [38]; [3] Taken from Cuijpers et al. 2019 [10]; [4] Taken from Cipriani et al. 2018 [39]; [5] Taken from Vittengl et al. 2007 [40]; [6] Taken from Biesheuvel-Leliefeld et al. 2015 [41].

model was refined using feedback by healthcare users, healthcare professionals, and researchers in the field of depression. DepMod adds to other health-economic models described in literature (e.g. [42–45]) by (i) considering both the prevention and treatment of depression, (ii) considering both psychological and pharmacological interventions, (iii) modeling increasing recurrence rates for patients with a higher number of previous episodes, (iv) considering a five-year time horizon, and (v) allowing for assessing the cost-effectiveness of interventions aiming to increase adherence. The model can be adapted to other settings by adding or removing interventions, by altering their coverage rate, adherence rate, or healthcare costs and by adjusting the epidemiology.

Outcomes of the model were shown in an example for the Dutch situation where the cost-effectiveness and budget impact were assessed of increasing the coverage of preventive interventions. Results of that analysis showed favorable outcomes suggesting that, when assuming a willingness-to-pay threshold of €20,000 per QALY, scaling up prevention is likely to be cost-effective. This finding is in line with previous research in this field that demonstrated that depression prevention is associated with probabilities in the range of 68–90% of being cost-effective [46–49]. This example demonstrates the ability of the model to assess the cost-effectiveness of alternative healthcare systems, while making use of an extensive evidence-base from epidemiology, clinical effectiveness and economic costs.

4.1. Strengths and limitations

The strength of our model is that it synthesizes information from different domains into a single model that can be used by policymakers as a decision-support tool for estimating the cost-effectiveness and budget impact of alternative intervention packages by incorporating fully tweakable scenarios. However, our model has a number of limitations that need to be acknowledged.

First of all, because information from different domains was combined, it was necessary to make assumptions, for example, on how effectiveness of interventions impacted on epidemiology, thereby introducing uncertainty in the model's outcomes. Whenever assumptions had to be made, we favored conservative assumptions that are more likely to underestimate the cost-effectiveness of interventions.

A second limitation of DepMod, common in health-economic models, is the presence of uncertainty regarding both input parameters like transition probabilities (parameter uncertainty) and structural model choices like the chosen health states or type of analytical framework (model uncertainty). Parameter uncertainty was addressed using probabilistic sensitivity analysis on the parameters of interest. Although we partially account for model uncertainty, the main value of our model lies in interpreting the comparison between the outcomes of two scenarios, rather than the absolute costs and effects associated with a single scenario.

A third limitation is that there is no unique solution when determining each of the calibration parameters of the Markov model. Given the episodic nature of depressive disorder, the epidemiology cannot be modeled using just incidence, prevalence, and excess mortality, as would be the case with chronic conditions. Instead, information is required on remission rates as well as recurrence rates, adding degrees of freedom which takes away the possibility of a unique set of parameters defining an internally consistent epidemiology. Calibration parameters were set to closely match available epidemiological evidence. The model's resulting epidemiology was in line with the epidemiological evidence-base, with the exception of two of 13 recurrence rates, which were below the confidence interval as reported by Solomon et al. (2000); thus indicating that the risk of recurrence in DepMod is slightly lower in these instances, thereby resulting in conservative estimates of the cost-effectiveness of preventive interventions [32]. Final parameter choices were checked for face-validity with experts in the field of depression. Future research should aim to validate these findings.

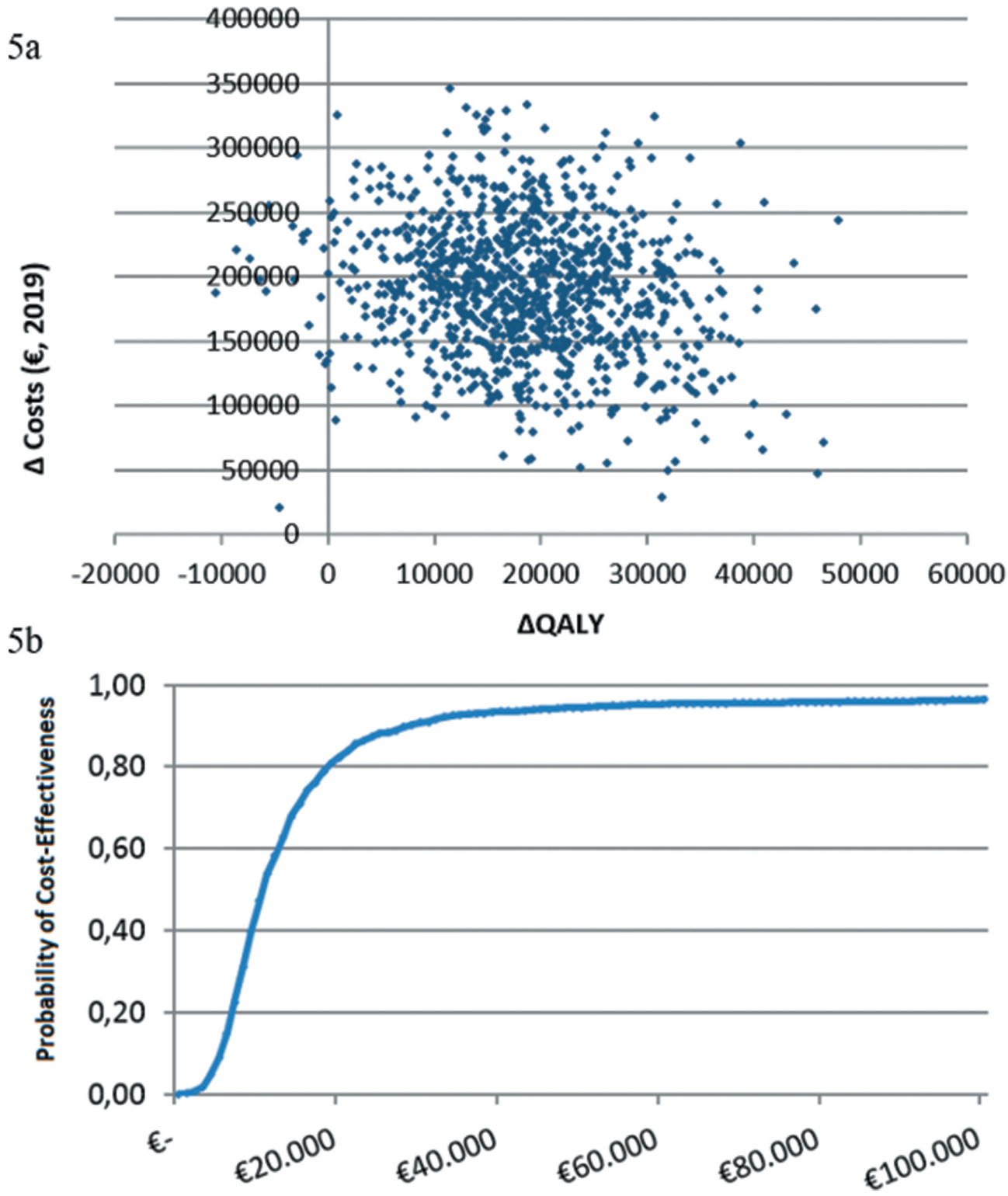

Figure 5. (a) Cost-effectiveness plane (top) and (b) cost-effectiveness acceptability curve (bottom) associated with scaling up prevention. The cost-effectiveness acceptability curve expresses the probability that scaling up prevention is cost-effective (y-axis) and on the x-axis the willingness to pay for one QALY gained given various ceiling ratios.

Fourth, we chose to model epidemiology using a Markov-model rather than a discrete event model. Both models have their strengths and limitations. As our aim was to make a population-level model, we were less interested in the ability of our model to capture individual traits affecting the course of illness. Nevertheless, cost-effectiveness might be affected by clinicians adjusting their treatment decisions based on the patient's personal traits. We partly compensated for this by modeling two forms of heterogeneity in the epidemiology. First, the severity of depression (mild, moderate, and severe)

was captured in the model. Second, the number of previous episodes was accounted for in the model using tunnel states. By adding these features, patient heterogeneity was considered to a limited extent. Moreover, as nuances such as medication switching were not specifically incorporated, ideally they should be captured within the parameters of each intervention (e.g., included in the average costs and average effect size of pharmacotherapy).

Fifth, we used one-year cycles in the model as available epidemiological evidence did not provide enough information for a shorter cycle-length. This allowed only one transition or episode of depression per year, even though changes in the course of depression can occur more frequently.

Sixth, ideally all healthcare costs, not only in the mental healthcare services, would be used for each health state, as depression could be associated with additional healthcare usage in general. Moreover, one would ideally also have data on the care patients receive when in intermediate (remissions and recurrences) states. These types of costs, but also other types of costs such as productivity losses, are not covered in the model. Moreover, although the model is unable to track how many treatments each patient has received (i.e., given it is a population model) and in theory patients could receive the same treatment multiple times during the five-year time horizon (each time with the same assumed effectiveness), this is likely only an issue when users would implement a 100% coverage rate for a specific intervention. In addition, the model is not optimally suited for treatments with a treatment duration longer than the time horizon (i.e., 1 year), given that treatments are offered every cycle.

Seventh, using NEMESIS-2 as our main source of epidemiological input, DepMod is restricted to the population aged 18–65 years. Using the model for other age-groups requires updating the input parameters to the setting and age group of interest.

Eighth, people in DepMod remain in their initial mild, moderate, or severe depression severity level throughout the entire period under consideration, whereas in reality people could transition between severity levels. Also, our model is limited to major depressive disorder and does not consider dysthymia.

Ninth, DepMod reports on the desirability of an alternative setup of the healthcare system from a perspective of cost-effectiveness, but does not take into account other perspectives, such as equity, feasibility, sustainability, and acceptability.

Tenth, QALYs were calculated based on disability weights instead of utility values based on Stouthard et al. (1997) and Sanderon et al.'s (2004) conversion factor. A disability weight is a weight factor that reflects the severity of the disease on a scale from 0 (perfect health) to 1 (equivalent to death). In contrast, a utility represents the quality of a health state on a scale from 0 (equivalent to death) to 1 (perfect health). In general, a utility value incorporates a broader spectrum of dimensions of health state, not merely focusing on disease related factors and may also lead to states worse than death. However, the disability weights derived by Stouthard et al. (1997) were based on the EuroQol-5D instrument, an instrument commonly used to obtain utility values and therefore captures this broader view.

Last, DepMod constitutes a relatively short time horizon, in avoid extrapolating too far from the available epidemiological evidence-base. However, the implementation of a life time horizon may enable decision makers to more adequately capture societal costs, for example, it is known that depression is frequently associated with low marital quality, low work performance, and low earnings, thereby increasing societal costs and decreasing quality of life [50]. The episodic nature of depression underscores the importance of identifying individuals at ultra-high risk of relapse or recurrence, as enduring depression symptoms may require further treatment and continue to generate personal, financial, and societal costs [51, 52]. Future research should therefore focus on developing a health-economic model using patient-level simulations (e.g., a discrete-event simulation) to determine the impact of personalized preventive actions targeting patients at high risk of a recurrent depression.

5. Conclusion

Overall, our model estimates the cost-effectiveness of changing the configuration of a healthcare system for depression. Outcomes of the model should always be considered in light of other equally important (societal) values such as equity, feasibility, sustainability, and acceptability (see for example [52]). Above all, our model aims to facilitate decision-making with respect to optimizing the healthcare system for major depressive disorder but should never be seen as an auto-pilot for decision-making.

Authors contributions

Overall project coordination: JL, FS. Writing of manuscript: JL, BW. Critical appraisal of manuscript: HGR, JS, AM, RS, MDB, PC, FS. All authors have read and approved the manuscript.

Acknowledgments

We acknowledge dr. Talitha Feenstra, prof. dr. Claudi Bockting, and the taskforce for the development of the Dutch Multidisciplinary Guideline for Depression for constructive comments on early drafts of this manuscript.

Availability of data and materials

The model (DepMod) is freely available upon request at the authors.

Declaration of interest

The authors have no other relevant affiliations or financial involvement with any organization or entity with a financial interest in or financial conflict with the subject matter or materials discussed in the manuscript apart from those disclosed.

Reviewers disclosure

Peer reviewers on this manuscript have no relevant financial relationships or otherwise to disclose.

Ethics approval and consent to participate

Not applicable, the study did not involve human subjects.

ORCID

Joran Lokkerbol http://orcid.org/0000-0001-9949-5442
Ben Wijnen http://orcid.org/0000-0001-7993-1905
Pim Cuijpers http://orcid.org/0000-0001-5497-2743

References

Papers of special note have been highlighted as either of interest (•) or of considerable interest (••) to readers.

1. Ferrari AJ, Charlson FJ, Norman RE, et al. Burden of depressive disorders by country, sex, age, and year: findings from the global burden of disease study 2010. PLoS Med. 2013;10(11):e1001547.
2. Murray CJ, Lopez AD, Organization WH. The global burden of disease: a comprehensive assessment of mortality and disability from diseases, injuries, and risk factors in 1990 and projected to 2020: summary. World Health Organization, 1996.
3. Murray CJ, Vos T, Lozano R, et al. 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. 2012;380(9859):2197–2223.
4. Üstün TB, Ayuso-Mateos JL, Chatterji S, et al. Global burden of depressive disorders in the year 2000. Br J Psychiatry. 2004;184(5):386–392.
5. Vos T, Flaxman AD, Naghavi M, et al. 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. 2012;380(9859):2163–2196.
6. Whiteford HA, Ferrari AJ, Degenhardt L, et al. The global burden of mental, neurological and substance use disorders: an analysis from the Global Burden of Disease Study 2010. PloS One. 2015;10(2):e0116820.
7. Smit F, Cuijpers P, Oostenbrink J, et al. Costs of nine common mental disorders: implications for curative and preventive psychiatry. J Ment Health Policy Econ. 2006;9(4):193-200.
8. Sobocki P, Jönsson B, Angst J, et al. Cost of depression in Europe. J Ment Health Policy Econ. 2006;9(2):87–98.
9. Cuijpers P, van Straten A, Andersson G, et al. Psychotherapy for depression in adults: a meta-analysis of comparative outcome studies. J Consult Clin Psychol. 2008;76(6):909.
10. Cuijpers P, van Straten A, van Schaik A, et al. Psychological treatment of depression in primary care: a meta-analysis. Br J Gen Pract. 2009;59(559):e51–e60.
11. Cuijpers P, Karyotaki E, de Wit L, et al. The effects of fifteen evidence-supported therapies for adult depression: a meta-analytic review. Psychother Res. 2020;30(3):279–293.
• **Meta-analysis demonstrating effectiveness of depression treatments**
12. Cipriani A, Salanti G, Furukawa TA, et al. Antidepressants might work for people with major depression: where do we go from here? Lancet Psychiatry. 2018;5(6):461–463.
13. Cuijpers P, Noma H, Karyotaki E, et al. A network meta-analysis of the effects of psychotherapies, pharmacotherapies and their combination in the treatment of adult depression. World Psychiatry. 2020;19(1):92–107.
14. Cuijpers P, van Straten A, Smit F, et al. Preventing the onset of depressive disorders: a meta-analytic review of psychological interventions. Am J Psychiatry. 2008;165(10):1272–1280.
15. van Zoonen K, Buntrock C, Ebert DD, et al. Preventing the onset of major depressive disorder: a meta-analytic review of psychological interventions. Int J Epidemiol. 2014 Apr;43(2):318–329.
16. Vos T, Carter R, Barendregt J, et al. Assessing cost-effectiveness in prevention: ACE–prevention September 2010 final report. University of Queensland; 2010.
17. Chisholm D. Choosing cost-effective interventions in psychiatry: results from the CHOICE programme of the World Health Organization. World Psychiatry. 2005;4(1):37.
18. Drummond MF, Sculpher MJ, Claxton K, et al. Methods for the economic evaluation of health care programmes. Oxford, UK: Oxford university press; 2015.
19. Roberts M, Russell LB, Paltiel AD, et al. Conceptualizing a model: a report of the ISPOR-SMDM modeling good research practices task force–2. Med Decis Mak. 2012;32(5):678–689.
20. GGZ guidelines. Multidisciplinary guideline for depression, second revision. Utrecht: Trimbos Institute; 2011.
21. Bijl R, Van Zessen G, Ravelli A, et al. The Netherlands mental health survey and incidence study (NEMESIS): objectives and design. Soc Psychiatry Psychiatr Epidemiol. 1998;33(12):581–586.
22. Bijl RV, de Graaf R, Ravelli A, et al. Gender and age-specific first incidence of DSM-III-R psychiatric disorders in the general population. Soc Psychiatry Psychiatr Epidemiol. 2002;37(8):372–379.
23. de Graaf R, Ten Have M, Tuithof M, et al. First-incidence of DSM-IV mood, anxiety and substance use disorders and its determinants: results from the Netherlands Mental Health Survey and Incidence Study-2. J Affect Disord. 2013;149(1–3):100–107.
•• **Considerable interest: important parameter(s) in the model**
24. Spijker J, de Graaf R, Bijl RV, et al. Duration of major depressive episodes in the general population: results from The Netherlands Mental Health Survey and Incidence Study (NEMESIS). Br J Psychiatry. 2002 Sep;181:208–213.
25. Spijker J, De Graaf R, Bijl R, et al. Functional disability and depression in the general population. Results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS. Acta Psychiatr Scand. 2004;110(3):208–214.
26. Vos T, Haby MM, Barendregt JJ, et al. The burden of major depression avoidable by longer-term treatment strategies. Arch Gen Psychiatry. 2004;61(11):1097–1103.
27. Stouthard ME, Essink-Bot M, Bonsel G, et al. Disability weights for diseases in the Netherlands. Department of Public Health, Erasmus University Rotterdam, the Netherlands. 1997.
•• **Considerable interest: important parameter(s) in the model**
28. Sanderson K, Andrews G, Corry J, et al. Using the effect size to model change in preference values from descriptive health status. Qual Life Res. 2004;13(7):1255–1264.
•• **Considerable interest: important parameter(s) in the model**
29. Hakkaart-van Roijen L, Van der Linden N, Bouwmans C, et al. Kostenhandleiding. In: Methodologie van kostenonderzoek en referentieprijzen voor economische evaluaties in de gezondheidszorg In opdracht van Zorginstituut Nederland Geactualiseerde versie. Zorginstituut Nederland, Diemen, the Netherlands. 2015.
30. Van Hout BA, Al MJ, Gordon GS, et al. Costs, effects and C/E-ratios alongside a clinical trial. Health Econ. 1994;3(5):309–319.
31. Naimark DM, Bott M, Krahn M. The half-cycle correction explained: two alternative pedagogical approaches. Med Decis Mak. 2008;28(5):706–712.
32. Solomon DA, Keller MB, Leon AC, et al. Multiple recurrences of major depressive disorder. Am J Psychiatry. 2000 Feb;157(2):229–233.
•• **Considerable interest: important parameter(s) in the model**
33. Briggs A, Sculpher M, Claxton K. Decision modelling for health economic evaluation. Oxford, UK: OUP Oxford; 2006.
34. de Graaf R, Ten Have M, van Gool C, et al. Prevalence of mental disorders and trends from 1996 to 2009. Results from the Netherlands Mental Health Survey and Incidence Study-2. Soc Psychiatry Psychiatr Epidemiol. 2012;47(2):203–213.
35. Chisholm D, Sanderson K, Ayuso-Mateos JL, et al. Reducing the global burden of depression: population-level analysis of intervention cost-effectiveness in 14 world regions. Br J Psychiatry. 2004;184(5):393–403.
36. Voor de Statistiek CB. Levensverwachting; geslacht, leeftijd (per jaar en periode van vijf jaren). 2018. Available from: https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37360ned/table?

37. Lokkerbol J, Adema D, Cuijpers P, et al. Improving the cost-effectiveness of a healthcare system for depressive disorders by implementing telemedicine: a health economic modeling study. Am J Geriatric Psychiatry. 2014;22(3):253–262.
38. Karyotaki E, Riper H, Twisk J, et al. Efficacy of self-guided internet-based cognitive behavioral therapy in the treatment of depressive symptoms: a meta-analysis of individual participant data. JAMA Psychiatry. 2017;74(4):351–359.
• **Used to derrive estimates for illustrative example**
39. Cipriani A, Furukawa TA, Salanti G, et al. Comparative efficacy and acceptability of antidepressant drugs in the acute treatment of major depressive disorder: a network meta-analysis. Lancet. 2018;391:1357–1366.
• **Used to derrive estimates for illustrative example**
40. Vittengl JR, Clark LA, Dunn TW, et al. Reducing relapse and recurrence in unipolar depression: a comparative meta-analysis of cognitive-behavioral therapy's effects. J Consult Clin Psychol. 2007;75(3):475.
• **Used to derrive estimates for illustrative example**
41. Biesheuvel-Leliefeld KE, Kok GD, Bockting CL, et al. Effectiveness of psychological interventions in preventing recurrence of depressive disorder: meta-analysis and meta-regression. J Affect Disord. 2015;174:400–410.
• **Used to derrive estimates for illustrative example**
42. Doyle JJ, Casciano J, Arikian S, et al. A multinational pharmacoeconomic evaluation of acute major depressive disorder (MDD): a comparison of cost-effectiveness between venlafaxine, SSRIs and TCAs. Value Health. 2001;4(1):16–31.
43. Nuijten MJ, Hardens M, Souêtre EA. Markov process analysis comparing the cost effectiveness of maintenance therapy with citalopram versus standard therapy in major depression. Pharmacoeconomics. 1995;8(2):159–168.
44. Paulden M, Palmer S, Hewitt C, et al. Screening for postnatal depression in primary care: cost effectiveness analysis. BMJ. 2009;339:b5203.
45. Simpson KN, Welch MJ, Kozel FA, et al. Cost-effectiveness of transcranial magnetic stimulation in the treatment of major depression: a health economics analysis. Adv Ther. 2009;26(3):346–368.
46. Wells KB, Schoenbaum M, Duan N, et al. Cost-effectiveness of quality improvement programs for patients with subthreshold depression or depressive disorder. Psychiatric Serv. 2007;58(10):1269–1278.
47. van der Aa HP, van Rens GH, Bosmans JE, et al. Economic evaluation of stepped-care versus usual care for depression and anxiety in older adults with vision impairment: randomized controlled trial. BMC Psychiatry. 2017;17(1):280.
48. Lynch FL, Hornbrook M, Clarke GN, et al. Cost-effectiveness of an intervention to prevent depression in at-risk teens. Arch Gen Psychiatry. 2005;62(11):1241–1248.
49. Fernández A, Mendive JM, Conejo-Cerón S, et al. A personalized intervention to prevent depression in primary care: cost-effectiveness study nested into a clustered randomized trial. BMC Med. 2018;16(1):28.
50. Kessler RC. The costs of depression. Psychiatr Clinics. 2012;35(1):1–14.
51. Wojnarowski C, Firth N, Finegan M, et al. Predictors of depression relapse and recurrence after cognitive behavioural therapy: a systematic review and meta-analysis. Behav Cogn Psychother. 2019;47(5):514–529.
52. Mihalopoulos C, Magnus A, Lal A, et al. Is implementation of the 2013 Australian treatment guidelines for posttraumatic stress disorder cost-effective compared to current practice? A cost-utility analysis using QALYs and DALYs. Aust N Z J Psychiatry. 2015;49(4):360–376.
• **Used to derrive estimates for illustrative example**

Appendix 1

Tabulated results of real-world example

	Total costs (95%CI)	Total QALYs (95%CI)	Incremental costs	Incremental QALYs	ICER
Reference scenario	€632,198 (€566,008 – €700,627)	36,662 (27,079–48,500)	NA	NA	NA
Alternative scenario	€823,304 (€748,926 – €898,681)	(55,069 (42,569–71,031)	€191,106	18,408	€10 per QALY gained