CrossMark

Model-Based Economic Evaluation of Treatments for Depression: A Systematic Literature Review

Spyros Kolovos[1] · Judith E. Bosmans[1] · Heleen Riper[2] · Karine Chevreul[3,4,5] · Veerle M. H. Coupé[6] · Maurits W. van Tulder[1]

Abstract

Background An increasing number of model-based studies that evaluate the cost effectiveness of treatments for depression are being published. These studies have different characteristics and use different simulation methods.

Objective We aimed to systematically review model-based studies evaluating the cost effectiveness of treatments for depression and examine which modelling technique is most appropriate for simulating the natural course of depression.

Methods The literature search was conducted in the databases PubMed, EMBASE and PsycInfo between 1 January 2002 and 1 October 2016. Studies were eligible if they used a health economic model with quality-adjusted life-years or disability-adjusted life-years as an outcome measure. Data related to various methodological characteristics were extracted from the included studies. The available modelling techniques were evaluated based on 11 predefined criteria.

Results This methodological review included 41 model-based studies, of which 21 used decision trees (DTs), 15 used cohort-based state-transition Markov models (CMMs), two used individual-based state-transition models (ISMs), and three used discrete-event simulation (DES) models. Just over half of the studies (54%) evaluated antidepressants compared with a control condition. The data sources, time horizons, cycle lengths, perspectives adopted and number of health states/events all varied widely between the included studies. DTs scored positively in four of the 11 criteria, CMMs in five, ISMs in six, and DES models in seven.

Conclusion There were substantial methodological differences between the studies. Since the individual history of each patient is important for the prognosis of depression, DES and ISM simulation methods may be more appropriate than the others for a pragmatic representation of the course of depression. However, direct comparisons between the available modelling techniques are necessary to yield firm conclusions.

✉ Spyros Kolovos
s.kolovos@vu.nl

[1] Department of Health Sciences, Faculty of Earth and Life Sciences, EMGO+ Institute for Health and Care Research, VU University Amsterdam, De Boelelaan 1085, 1081 HV Amsterdam, The Netherlands

[2] Department of Clinical, Neuro and Developmental Psychology, Faculty of Behavioural and Movement Sciences, EMGO+ Institute for Health and Care Research, VU University Amsterdam, Amsterdam, The Netherlands

[3] URC Eco Ile de France, AP-HP, Paris, France

[4] Université Paris Diderot, Sorbonne Paris Cité, ECEVE, UMRS 1123, Paris, France

[5] INSERM, ECEVE, U1123, Paris, France

[6] Department of Epidemiology and Biostatistics, VU University Medical Center, Amsterdam, The Netherlands

Key Points for Decision Makers

Substantial methodological differences exist between model-based studies evaluating depression treatments.

Simpler models, such as decision trees and cohort-based Markov models, were more frequently used.

Microsimulation models, such as individual-based state-transition models and discrete-event simulation models, incorporate patient heterogeneity and history, which is important when modelling depression.

1 Introduction

Depression is associated with an enormous burden for patients, healthcare systems and society as a whole. The World Health Organization has estimated that depression will rank first among the most debilitating conditions in terms of disability-adjusted life-years (DALYs) by the year 2030 [1]. Individuals with depression experience impairments in many aspects of their life, resulting in significantly lower quality of life (QoL) than the general population [2–4]. Additionally, depression leads to an immense economic burden [5]. In the USA, the total costs of depression were estimated to be $US173 billion in 2005 and $US202 billion in 2010 [6], whereas in Europe the total expenses related to depression were estimated to be approximately €92 billion in 2010 [7, 8].

Depression is very common, with a lifetime prevalence between 10 and 15% [9]. Patients with depression constitute a rather heterogeneous group [10], which may also affect the effectiveness and cost effectiveness of treatments [11, 12]. Evidence exists that various evidence-based treatments are similarly effective against depression, but whether specific treatments can be more beneficial for patients with specific characteristics remains unclear [13]. More than 50% of the patients seeking treatment for depression experience recurrent depressive episodes, the mean duration of which is 16 weeks [14]. Previous episodes are significantly related to the probability of recurrence, with the risk for a new depressive episode increasing by 16% for each successive recurrence [15]. Furthermore, the probability of recurrence decreases as the period that the patient is in recovery increases. Depression often is comorbid with other mental disorders, such as anxiety or substance abuse, or physical disorders such as diabetes [14, 16]. Therefore, the personal history of each individual is important for the prognosis of depression.

Economic evaluations are frequently used to inform decisions about the allocation of scarce resources in healthcare [17, 18]. They can be conducted either alongside randomized controlled trials (RCTs) or using economic modelling techniques. The current paper focuses on modelling techniques. The three dominant modelling methods used in model-based studies for common mental health disorders are decision trees (DTs), state-transition Markov models (MMs), and discrete-event simulation (DES) models [19–21]. These modelling techniques have different strengths and limitations, and the choice of one over the other depends on several factors, such as the complexity of the decision problem, data accessibility and available expertise [22]. A recent literature review compared MMs and DES for the economic evaluation of healthcare technologies such as breast cancer treatments without focusing specifically on depression [20]. The authors concluded that DES showed a few profound advantages over MMs but that the final choice depends on the research question being addressed [20]. So far, the most appropriate modelling method to use when simulating the course of depression and the impact of treatment on this course remains unclear.

Economic modelling in the research field of depression has received increased attention in recent years. Afzali et al. [21] conducted a systematic literature search up to September 2010 reviewing model-based cost-effectiveness studies of depression treatments. The authors focused on the characteristics of the model-based studies and evaluated how these characteristics influenced the results. They identified wide variation in the methodological aspects of the included studies and concluded that these differences potentially influence the accuracy of the cost-effectiveness estimations [21]. Another systematic review examined model-based studies of pharmacotherapy for major depression, published up to May 2010, and reached similar conclusions [23].

Many additional modelling studies have been published since 2010, and we considered it opportune to update the review by Afzali et al. [21] because periodic updating of evidence is important to avoid stakeholders using outdated evidence to inform decisions or address new problems. Therefore, the first objective of the current study was to investigate the methodological characteristics of model-based studies examining the cost effectiveness of treatments for depression. Second, we aimed to examine which of the modelling methods was more appropriate for simulating the course of depression.

2 Methods

2.1 Literature Search

To answer the first research question, we updated the systematic review by Afzali et al. [21]. The first author

(SK) searched the databases PubMed, EMBASE and PsycInfo for eligible studies published between 1 January 2002 and 1 October 2016. This timeframe overlaps with that examined by Afzali et al. [21] (2002–2010) because we developed our own search strategy and wanted to ensure the same search criteria were applied for all years considered. We used search terms referring to model-based economic evaluations and depression treatments (see Appendix 1 in the Electronic Supplementary Material [ESM] for the full search string). We searched for published studies written in English. No protocol was published for this review. We input all entries from the database searches into a reference manager and removed the duplicates. The titles and abstracts were screened for eligibility, and the full texts were then examined.

2.2 Inclusion Criteria

Studies were included when they fulfilled the following criteria: (1) used a health economic model, (2) examined the cost effectiveness of treatments for adults with depression and (3) estimated quality-adjusted life-years (QALYs) or DALYs. Health economic models were defined as a mathematical representation of reality that can be used to estimate the cost effectiveness of health technologies for depression [21]. Depression was defined as a diagnosis of major depressive disorder based on a structured clinical interview or a score on a standardized self-report measure of depressive symptom severity that indicates the presence of clinically relevant depression. We included studies using QALYs or DALYs, which is similar to the inclusion criteria used by Afzali et al. [21]. This approach was adopted to increase the homogeneity of the reviewed studies. At the same time, the amelioration of the QoL of patients with depression is one of the main goals of treatment for depression and thus a relevant outcome in the context of this study [24]. No restrictions were applied on treatments evaluated (i.e. pharmacotherapy, psychotherapy, combination treatments, other) or the control groups. Studies on the prevention of depression were excluded. Any ambiguity around whether a study should be included was discussed with another author (JEB) until consensus was reached.

2.3 Data Extraction and Analysis

Data were extracted on the modelling technique used (e.g. DTs, cohort-based state-transition MMs [CMMs]), the structure of the model (i.e. what health states/events were used to represent the natural course of depression), data sources (e.g. meta-analysis, RCTs), time horizon and cycle length, treatment comparisons (i.e. which intervention and control groups were compared) and sensitivity analyses and

economic perspective (e.g. societal or healthcare system perspective). Data extraction was performed by the first author (SK) using a data form developed based on the review by Afzali et al. [21]. When data were missing or unclear, we emailed the first author of the paper to request additional information. The extracted data were presented using descriptive statistics such as means and percentages whenever possible.

2.4 Comparison of Modelling Techniques

We grouped the modelling methods into four classes: DTs, CMMs, individual-based state-transition models (ISMs), and DES models (see the Results section for a description of each modelling method). We used 11 predefined criteria to evaluate the appropriateness of each modelling technique for simulating the course of depression. These criteria were introduced by Brennan et al. [25] to provide guidance on key requirements of modelling methods and have been previously used in schizophrenia [26]. We aimed to extract information on the included studies for each criterion (e.g. how much building time the model required, were data available as required to populate the model, did the authors comment on the simulation time required, etc.). If information was not available in the included studies for a specific criterion, we searched the literature to identify data on the performance of each model on this criterion. The criteria were building time, data collection, experience, simulation time, clinical representation, patient heterogeneity, timing of events, memory, patient interaction, interaction due to covariates and variability (see Table 1 for a short definition of each criterion). We discussed the performance of each model in each criterion specifically for depression.

3 Results

3.1 Literature Search

Figure 1 describes the results of the literature search. The search yielded 7757 titles; 6570 references remained after duplicates were removed. After screening the titles and abstracts, we retrieved the full texts of 113 studies and excluded 72 (16 no depression, 35 no model-based studies, 18 no QALYs or DALYs, two prevention of depression, and one no adults). We included 41 studies compared with the 14 studies included by Afzali et al. [21] (Fig. 1).

The included studies used different health economic modelling techniques (Table 2): 21 (51%) studies used DTs, 15 (37%) used CMMs, two (5%) used ISMs and three (7%) used DES modelling techniques.

Table 1 Description of 11 criteria used to evaluate the appropriateness of each modelling technique to simulate the course of depression. Adapted from Heeg et al. [26]

| Criterion	Description
Building time	Time required to develop and program the model
Data collection	Availability and time required to collect data to populate the model
Experience	Number of published studies using each modelling method in healthcare literature
Simulation time	Computational time needed to generate results
Clinical representation	Ability to simulate the problem pragmatically by including all the relevant aspects
Patient heterogeneity	Ability to deal with patients having different clinical and demographic characteristics
Timing of events	Ability to allow for recurrent events to occur
Memory	Ability to track each patient and allow the personal history to influence future events
Patient interaction	Ability to include interaction between patients
Interaction due to covariates	Ability to include covariates that interact or influence future events
Variability	Ability to incorporate and analyse uncertainty within the model structure
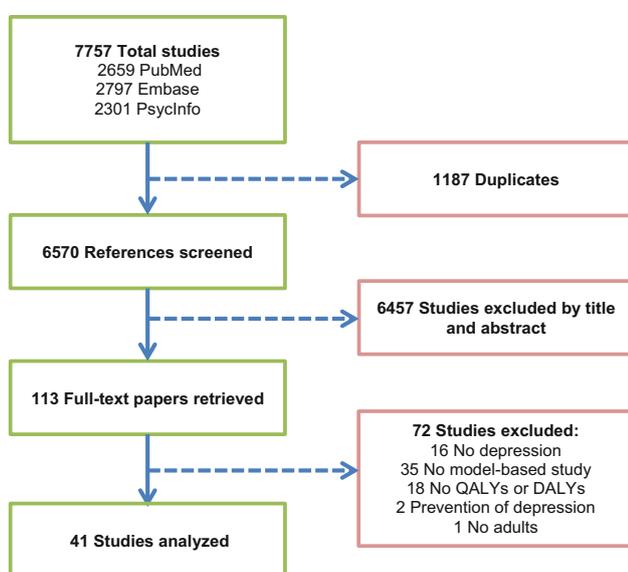
Fig. 1 Flowchart describing the search strategy. *DALY* disability-adjusted life-year, *QALY* quality-adjusted life-year

3.2 Characteristics of Model-Based Studies for Depression

3.2.1 Time Horizon and Cycle Length

In the studies using DTs, the time horizon varied considerably between 6 weeks and 27 months. Likewise, in those using CMMs, the time horizon varied between 4 months and lifetime and the cycle length ranged from 1 week to 6 months (Table 2). In the two studies using ISMs, the time horizon was 3 and 5 years and the cycle length 2 months and 1 month, respectively. Finally, one of the three studies that used DES models had a time horizon of 5 years [27] and the other two used a lifetime time horizon (Table 2) [28, 29].

3.2.2 Perspective

In total, 12 (29%) studies adopted the societal perspective, which included all relevant direct and indirect healthcare costs and lost productivity costs, and 21 studies (51%) used the healthcare perspective, which included only healthcare costs. Six studies (16%) used the managed care organization perspective, which included the costs from the perspective of the organization providing the healthcare (Table 2), and, finally, one study (2%) used the employer's perspective and one study (2%) used the payer's perspective.

3.2.3 Treatment and Comparator

Just over half of the studies (22 [54%]) evaluated the cost effectiveness of pharmacological treatments for depression. Seven (17%) examined combined treatments, mostly a combination of antidepressants and psychological treatments. Four (10%) investigated psychological treatments, and two (5%) investigated antidepressant treatment in combination with genotyping. Two studies (5%) examined electroconvulsive therapy (ECT) and two (5%) studied repetitive transcranial magnetic stimulation (rTMS). Finally, two studies (5%) examined the cost effectiveness of positive mental training and St John's Wort (Table 2).

Table 2 also presents the comparator in the included models: 23 (56%) studies used antidepressant medication as a comparator; seven (17%) used more than one comparator, including psychological treatments, no treatment, and combination treatments; and five (12%) compared the treatment under study with treatment as usual. Other comparators were computerized cognitive behavioural therapy, placebo pills, rTMS, and ECT (Table 2).

Table 2 Summary of model-based studies of depression

Study	Treatment	Comparators	Model type	Main events/health states	Time horizon	Cycle length	Main perspective
Armstrong et al. [68]	Escitalopram	Antidepressants	DT	Response, second-line options	6 months	NA	Managed care organization
Armstrong et al. [69]	Escitalopram	Duloxetine	CMM	Remission, second-line options	1 year	Weekly	Managed care organization
Aziz et al. [70]	Antidepressants	ECT	CMM	Full remission, partial remission, no response, death	Lifetime	6 months	Societal
Beil et al. [71]	IPT	CBT, antidepressants	CMM[a]	Depression, remission, chronic depression	3 years	6 months	Managed care organization
Benedict et al. [42]	Duloxetine	Antidepressants	CMM	Remission, response, relapse, discontinuation, second-line options	1 year	2 months	Healthcare
Evans-Lacko et al. [72]	Antidepressants	CBT, CBT + antidepressants	DT	Remission, response, no response	27 months	NA	Employer
Francois and Toumi [73]	Escitalopram	Antidepressants	DT	Remission, relapse, discontinuation, adverse events, second-line options	6 months	NA	Societal
Greenhalgh et al. [74]	ECT	Antidepressants	DT	Depression, maintenance therapy, continue treatment, relapse	1 year	NA	Healthcare
Kaltenthaler et al. [75]	CCBT	TAU	DT	Minimal, mild, moderate, severe depressive symptoms	18 months	NA	Healthcare
Khoo et al. [76]	Antidepressants	Antidepressants	DT	Remission, relapse, augmentation, second-line options	6 months	NA	Societal
Koeser et al. [77]	PosMT	CCBT	DT	Minimal, mild, moderate, severe depressive symptoms	5 months	NA	Healthcare
Koeser et al. [78]	Antidepressants	CBT, CBT + antidepressants	DT	Remission, response, no response, relapse	27 months	NA	Healthcare
Kozel et al. [79]	rTMS	ECT	DT	Response, no response, relapse	1 year	NA	Healthcare
Leelahanaj [80]	Aripiprazole	Placebo	DT	Remission, non-remission, discontinuation	6 weeks	NA	Healthcare
Leelahanaj [81]	Bupropion	Antidepressants	DT	Remission, non-remission, discontinuation, adverse events	3 months	NA	Healthcare
Lenox-Smith et al. [82]	Venlafaxine	Antidepressants	DT	Remission, relapse, response, second-line options	6 months	NA	Healthcare
Maniadakis et al. [83]	Agomelatine	Antidepressants	CMM	Healthy, depression, remission, death	2 years	1 month	Societal
Mencacci et al. [35]	Antidepressants	Antidepressants	DT	Remission, no remission, relapse, suicide attempt, death	1 year	NA	Healthcare
Naversnik and Mrhar [84]	Web-based service + TAU	TAU	CMM	Minimal, mild, moderate, severe depressive symptoms	1 year	6 months	Healthcare
Nguyen and Gordon [85]	rTMS	Antidepressants	ISM	Treatment, full remission, partial remission, relapse, ECT, augmentation of antidepressants	3 years	2 months	Healthcare
Nordstrom et al. [86]	Escitalopram	Venlafaxine and duloxetine	DT	Remission, relapse	6 months	NA	Societal

Table 2 continued

Study	Treatment	Comparators	Model type	Main events/health states	Time horizon	Cycle length	Main perspective
Nuijten et al. [36]	Escitalopram	Antidepressants	DT	Remission, no remission, sustained remission, relapse, discontinuation, dose increase	6 months	NA	Societal
Olgiati et al. [43]	Antidepressants + genotyping	Antidepressants	CMM	Remission, no remission, drop-out	4 months	8 weeks	Managed care organization
Olgiati et al. [87]	Antidepressants	Antidepressants	CMM	Remission, no remission, drop-out	6 months	6.5 weeks	Healthcare
Prukkanone et al. [27]	Antidepressants and CBT	No treatment	DES	Remission/recovery, suicide, relapse/recurrence	5 years	NA	Healthcare
Ramsberg et al. [88]	Antidepressants	Venlafaxine	DT	Remission, no remission, suicide attempt, death, survival	1 year	NA	Societal
Sado et al. [89]	Combination treatment	Antidepressants	DT	Remission, relapse, discontinuation	1 year	NA	Societal
Saylan et al. [28]	Aripiprazole	Antidepressants	DES	Depressive episode, remission, death between episodes	Lifetime	NA	Payer
Serretti et al. [90]	Antidepressants + genotyping	Antidepressants	CMM	Remission, no remission, drop-out	3 months	8 weeks	Healthcare
Simon et al. [91]	Combination treatment	Antidepressants	DT	Remission, relapse, discontinuation	15 months	NA	Healthcare
Sisiking et al. [92]	Enhanced IPT	IPT, no intervention	CMM	No depression, depression, remission, chronic depression, death	Lifetime	6 months	Healthcare
Siskind et al. [44]	Improved stepped care	TAU, no treatment	CMM[a]	Depression, remission, chronic depression, dead	Lifetime	6 months	Healthcare
Sobocki et al. [93]	Antidepressants	TAU	ISM	Depressive episode, remission, well, dead	5 years	1 month	Societal
Sobocki et al. [94]	Venlafaxine	No treatment	CMM	Depressive episode, remission, well, dead	2 years	1 month	Societal
Solomon et al. [95]	St John's Wort	Antidepressants	CMM	Depressive episode, response, remission, dead	18 months	6 weeks	Healthcare
Solomon et al. [96]	CCBT	CBT, TAU	DT	Remission, no remission, maintenance	6 months	NA	Healthcare
Sullivan et al. [97]	Escitalopram	Antidepressants	DT	Response, adverse events, second-line options	9 months	NA	Managed care organization
Tosh et al. [29]	Prevent drop-out, self-referral, no therapy	TAU	DES	Normal mood, response, remission, recovery, depressive episode, relapse, recurrence	Lifetime	NA	Healthcare
Trivedi et al. [37]	Venlafaxine	Antidepressants	DT	Remission, relapse, response, discontinuation	2 months	NA	Managed care organization
Vallejo-Torres et al. [98]	ECT	rTMS, rTMS followed by ECT	CMM[a]	Response, remission, relapse, recurrence	1 year	2 weeks	Healthcare

Table 2 continued							
Study	Treatment	Comparators	Model type	Main events/health states	Time horizon	Cycle length	Main perspective
Wang et al. [45]	Enhanced care	TAU	CMM	Never depressed, depressed, recovered, dead	5 years	3 months	Societal

CBT cognitive behavioural therapy, *CCBT* computerized cognitive behavioural therapy, *CMM* cohort-based state-transition Markov model, *DES* discrete-event simulation, *DT* decision tree, *ECT* electroconvulsive therapy, *IPT* interpersonal therapy, *ISM* individual-based state-transition model, *NA* not applicable, *PosMT* positive mental thinking, *rTMS* transcranial magnetic stimulation, *TAU* treatment as usual

^a CMM used 'tunnel states' as a function of time (semi-Markov models)

3.2.4 Structure of the Models

The structure of the models varied considerably between the studies. The most frequently used health states or events were response, remission, relapse and death (Table 2). Three studies (7%) used different states for different levels of depression severity (i.e. minimal, mild, moderate and severe depressive symptoms), and 13 studies (32%) included one or more of the health states discontinuation, second-line options and adverse events related to antidepressant medication.

3.2.5 Data Sources

The data sources that can be used as inputs for a health economic model have different strengths and limitations, but they must all be transparent and well documented [30, 31]. The included studies used various data sources to populate the model (Table 3). These sources included but were not limited to published RCTs, meta-analyses of RCTs, observational studies and expert opinions.

A total of 27 studies (66%) used meta-analysis to derive clinical inputs. Since meta-analyses are based on a systematic literature search using predefined inclusion criteria, the possibility of selection bias is decreased. Expert opinion, which is considered an unreliable data source because of its subjective nature, was used in 15 studies (37%). Most of the studies used published data sources to estimate utility weights, such as RCTs or observational studies. Two studies (5%) used a depressive symptom severity measure to calculate utility weights, and two other studies (5%) converted depression-free days to utility weights (Table 3).

3.2.6 Sensitivity Analysis

All included studies carried out sensitivity analyses to examine the uncertainty around the study parameters (Table 3). The two basic purposes of sensitivity analysis are to investigate the robustness of a chosen course of action and to accumulate additional information surrounding a specific decision [32]. Of the included studies, 29 (70%) performed a univariate sensitivity analysis, six (15%) carried out a multivariate sensitivity analysis and six (15%) conducted both univariate and multivariate analyses (Table 3). The sensitivity analysis was probabilistic in 31 (77%) of the studies, deterministic in three (7%) and both probabilistic and deterministic in seven (17%).

3.3 Comparison of Modelling Techniques

We used 11 criteria to evaluate the appropriateness of the four modelling methods (DTs, CMMs, ISMs and DES) used in the included studies [25] and used these data to

Table 3 Data sources and sensitivity analysis

Study	Primary source of data			Sensitivity analysis
	Clinical inputs	Cost	Utility	
Armstrong et al. [68]	RCTs, claims database	Published literature, USA Medicare	Published longitudinal surveys using EQ-5D	Univariate sensitivity analysis, PSA
Armstrong et al. [69]	RCTs, claims database	Published literature, USA Medicare	Published longitudinal surveys using EQ-5D	Univariate sensitivity analysis, PSA
Aziz et al. [70]	RCTs, observational studies, national life tables	Published national data, published literature	Published literature	Univariate and multivariate sensitivity analysis, PSA
Beil et al. [71]	RCTs, observational studies, published literature	Unpublished clinical trial data, published database	Published literature	Univariate sensitivity analysis, DSA and PSA
Benedict et al. [42]	Meta-analysis of RCTs, expert opinion, published literature	Published national data, expert opinion	Unpublished trials using EQ-5D	Univariate sensitivity analysis, PSA
Evans-Lacko et al. [72]	National observational studies, RCTs, meta-analysis	Published literature, published database	Published literature	Univariate sensitivity analysis, DSA
Francois and Toumi [73]	Clinical trial, observational studies, expert opinion	Published national data, expert opinion	Unpublished literature	Univariate sensitivity analysis, break-even analysis, PSA
Greenhalgh et al. [74]	RCTs, literature review	Published national data	Published literature	Univariate sensitivity analysis, PSA
Kaltenthaler et al. [75]	RCTs, observational study	RCT, published national data,	RCTs, published national data, expert opinion	Univariate and multivariate sensitivity analysis, PSA
Khoo et al. [76]	Meta-analysis, published literature	Published literature, expert opinion	Published literature	Univariate sensitivity analysis, DSA and PSA
Koeser et al. [77]	Meta-analysis of RCTs	Published national data, published literature, RCT	Published literature	Univariate sensitivity analysis, DSA and PSA
Koeser et al. [78]	RCTs, meta-analysis, expert opinion	Published literature, published clinical trial data	Published clinical trial data	Univariate and multivariate sensitivity analysis, DSA and PSA
Kozel et al. [79]	Observational studies	Published national data, expert opinion	Published literature	Univariate sensitivity analysis, PSA
Leelahanaj [80]	Meta-analysis of RCTs	National sources	Published literature	Univariate sensitivity analysis, PSA
Leelahanaj [81]	Observational studies	National sources, published literature	Published literature	Univariate sensitivity analysis, PSA
Lenox-Smith et al. [82]	Meta-analysis of RCTs, expert opinion	UK NHS	Using a published approach to convert DFDs to utility weights	Univariate sensitivity analysis, rank order stability assessment, PSA
Maniadakis et al. [83]	Meta-analysis of RCTs	National sources, expert opinion	Published literature	Univariate sensitivity analysis, DSA and PSA
Mencacci et al. [35]	Meta-analysis, observational studies	National data, expert opinion, published literature	Published literature	Multivariate sensitivity analysis, DSA and PSA
Naversnik and Mrhar [84]	Meta-analysis	Published literature, expert opinion, RCT	Convert BDI scores to utility weights	Multivariate sensitivity analysis, PSA
Nguyen and Gordon [85]	Observational studies	Expert opinion, assumption	Published literature	Univariate and multivariate sensitivity analysis, PSA
Nordstrom et al. [86]	Meta-analysis of RCTs, clinical database	Published longitudinal survey, expert opinion, national data	Published longitudinal surveys using EQ-5D	Univariate sensitivity analysis, PSA
Nuijten et al. [36]	Meta-analysis of RCTs	Published national data	Published literature	Univariate sensitivity analysis, DSA
Olgiati et al. [43]	Observational studies, meta-analysis, published database	Published literature, observational studies	Published literature	Univariate and multivariate sensitivity analysis, PSA

Table 3 continued

Study	Primary source of data			Sensitivity analysis
	Clinical inputs	Cost	Utility	
Olgiati et al. [87]	Observational studies, meta-analysis	Published data, published literature	Published literature	Univariate sensitivity analysis, PSA
Prukkanone et al. [27]	Meta-analysis of RCTs	Published literature, national sources	Transformed effect sizes into DALYs	Multivariate sensitivity analysis, PSA
Ramsberg et al. [88]	Meta-analysis of RCTs	Published literature	Published literature	Multivariate sensitivity analysis, PSA
Sado et al. [89]	Meta-analysis of RCTs, clinical trial data, expert opinion	Published data from Japanese government	Published literature using SG	Univariate sensitivity analysis, PSA
Saylan et al. [28]	Meta-analysis, RCTs, clinical trials	Published national data	Published literature	Univariate sensitivity analysis, PSA
Serretti et al. [90]	Observational studies, meta-analysis, published database	National data, published databases, published literature	Published literature	Univariate sensitivity analysis, PSA
Simon et al. [91]	Meta-analysis of RCTs	Published national data	Published literature using SG	Univariate sensitivity analysis, PSA
Sisiking et al. [92]	Observational studies, published literature	Expert opinion	RCTs, observational studies, published literature	Univariate sensitivity analysis, PSA
Siskind et al. [44]	Meta-analysis, observational studies	Published literature	Published literature, RCT	Univariate sensitivity analysis, DSA
Sobocki et al. [93]	Meta-analysis of RCTs, observational study, national life tables	Published longitudinal survey, published national data	Published longitudinal surveys using EQ-5D	Univariate sensitivity analysis, PSA
Sobocki et al. [94]	Meta-analysis of RCTs, observational study, national life tables	Published longitudinal survey, published national data	Published longitudinal surveys using EQ-5D	Univariate sensitivity analysis, PSA
Solomon et al. [95]	Meta-analysis, unpublished trials	Published national data, Australian Medicare	Published literature	Univariate sensitivity analysis, PSA
Solomon et al. [96]	RCTs, meta-analysis	Published database, observational study	Published literature	Univariate and multivariate sensitivity analysis, PSA
Sullivan et al. [97]	Observational studies, medical records	Claims database	Published longitudinal surveys using EQ-5D	Univariate sensitivity analysis, PSA
Tosh et al. [29]	RCT, observational study, published data, expert opinion	UK NHS	Published literature	Univariate sensitivity analysis, PSA
Trivedi et al. [37]	Meta-analysis of RCTs	National sources and claims database	Using a published approach to convert DFDs to utility weights	Multivariate sensitivity analysis, PSA
Vallejo-Torres et al. [98]	Meta-analysis, RCT, expert opinion	National sources, published data, expert opinion	Published literature	Multivariate sensitivity analysis, DSA and PSA
Wang et al. [45]	RCTs, observational studies, national life tables	Published national data, published literature	Published literature using SG	Univariate sensitivity analysis, PSA

BDI Beck Depression Inventory, *DALYs* disability-adjusted life-years, *DFD* depression-free days, *DSA* deterministic sensitivity analysis, *EQ-5D* EuroQol 5-Dimensions, *NHS* national healthcare service, *PSA* probabilistic sensitivity analysis, *RCT* randomized controlled trial, *SG* standard gamble

evaluate the models on the predefined criteria. An overview of this evaluation is presented in Table 4. A description of the criteria applied to each modelling method, with a specific focus on the importance of these criteria when modelling the cost effectiveness of depression treatments, follows.

3.3.1 Decision Trees

DTs scored positively in four of the 11 criteria (Table 4). DTs are used to represent relatively simple decision problems and are mainly appropriate for modelling interventions over short time horizons [22]. This is reflected in

Table 4 Strengths and weaknesses of four modelling techniques for simulating the course of depression

Criterion	Decision tree	Cohort-based state-transition Markov model	Individual-based state-transition model	DES
Building time	+	+	−	−
Data collection	+	+	−	−
Experience	+	+	+	−
Simulation time	+	+	−	−
Clinical representation	−	−	+	+
Patient heterogeneity	−	−	+	+
Time of events	−	+	+	+
Memory	−	−	+	+
Patient interaction	−	−	−	+
Interaction due to covariates	−	−	−	+
Variability	−	−	+	+

DES discrete-event simulation, + indicates modelling technique was evaluated positively for this criterion, − indicates modelling technique was evaluated negatively for this criterion

our finding that the average time horizon in the studies using DTs was 10 months. They are generally simple to build and analyse. A DT typically includes a small number of health states and has low data requirements. This was also reflected in the studies included in this review, as the majority used between two and four health states (Table 2). This may pose a threat to the clinical representation of depression in these models, since depression typically has an intermittent course. DT is probably the most popular modelling technique in simulating depression ($n = 21$ [51%]) because of the low requirements for time to develop, data inputs and expertise. For example, most of the included studies used a dedicated software such as TreeAge Pro [33] or other statistical software such as Microsoft® Excel and WinBUGS [34] to develop the DTs.

A DT typically focuses on the average patient. For instance, the included studies using DTs did not differentiate between patients with comorbidity or different age groups while simulating the course of depression. Furthermore, DT does not track the individual history of a patient, and passing of time is not included in the model. This means that DTs do not include variability in patient characteristics over time (e.g. increase the probability of a future recurrence after each additional depressive episode). This is a disadvantage of DTs, because depression is a chronic disorder for many patients, and it is important to model long-term outcomes that are influenced by previous events.

The time horizon used in the included DT studies varied between 6 weeks and 27 months. This may be a short period for a recurrent disorder such as depression, for which the average episode is 16 weeks and more than half of patients experience more than one episode [14]. However, a longer time horizon would substantially increase the

complexity of the model, making it difficult to handle. In addition, patient heterogeneity and interactions between variables are typically not included in DTs, but this can be implemented by running separate models for each subgroup or adding more health states to the model, respectively. However, this may also lead to an exponential increase in model size. These issues may decrease the face validity of DT models for depression. Only three of the reviewed studies that included DTs explicitly reported validation methods. Two used expert panels to validate the data sources and one used individual panellists to verify the model estimates [35–37].

3.3.2 Cohort-Based State-Transition Markov Model

The CMM scored positively in five of the 11 criteria (Table 4). The number of studies using CMMs (15 [37%]) indicated it was the second most popular technique to model depression. A CMM is most suitable for situations in which the decision problem can be represented in terms of health states and the population under study is a closed cohort [38]. Portions of the cohort simultaneously move from one health state to the other. The model is 'memoryless', a condition known as the 'Markovian property', meaning that future health states depend only on the current health state and not on the sequence of previous health states [38]. Therefore, the included studies using CMM did not account for patient history since the initiation of the simulation (e.g. time spent in previous health states). This is an essential drawback of using a CMM to simulate depression, because various patient characteristics (e.g. baseline depression severity, comorbidity, previous episodes, etc.) play an important role in prognosis and treatment decisions [39–41].

Nevertheless, additional health states or 'tunnel' states can be added to model patient history and heterogeneity or timing of events such as time spent in a depressive episode; these models are also called semi-Markov models. However, this may increase the complexity of the model considerably, resulting in a model that is rather difficult for analysts and researchers to manage and for decision makers and other stakeholders to interpret [38]. Only three of the 13 included studies used 'tunnel states'. Thus, most studies using CMM did not include the timing of events and patient heterogeneity in the models.

CMM is relatively easy to build, and most of the included studies used Microsoft® Excel and Visual Basic applications or dedicated software such as TreeAge Pro [33] to develop the model. An advantage of CMM is that, similar to DTs, it has relatively low data requirements. However, and in comparison with DTs, it is easier for CMM to simulate different courses of disease and thus it has better clinical representation and face validity. Other validation methods were reported in four CMM studies, including validation of the model structure and approach by a panel of clinicians, verification of the model algorithm (internal validity), validation of the model's assumptions by experts, and comparison of model predictions with estimates from published literature (external validity) [42–45].

3.3.3 Individual-Based State-Transition Model

The ISM scored positively in six of the 11 criteria (Table 4). Two studies employed an ISM that was a state-transition model such as CMM but was not limited by the Markovian property [38]. Instead of cohort simulation, ISM is a microsimulation model that simulates one patient at a time. The two included studies that used ISM tracked the individual history of the patients (in terms of increased risk for a future depressive episode after each new episode). They also included patient heterogeneity in the model by providing different transition probabilities based on patient characteristics, such as sex and age. In other words, ISM can include timing of events and memory in the model. However, this may increase the development and running time of the model considerably. The two studies used TreeAge Pro [33] to build the model.

The complexity of the model means data requirements are also higher than with DT and CMM (e.g. separate transition probabilities must be found for men and women or patients in different age groups), and this is a considerable issue. Although the volume of meta-analyses, RCTs and observational studies investigating depression in the literature is always increasing, there is no consensus on instruments to monitor the severity of depression and on the definitions of related health states such as remission and

response. As a result, searching for data to populate the models can be inefficient and time consuming. Data availability was also an issue in the included studies; 37% relied at least partly on expert opinion as a data source for some clinical or cost parameters.

A further issue related to the use of ISM that also applies to CMM is the need to define a fixed cycle length [38, 46]. The choice of cycle length depends upon treatment length, data availability and frequency of clinical events, among others. Shorter cycle lengths give a better approximation of the events in real life but increase the computational burden in terms of data requirements and running time [38]. The cycle length should be short enough to incorporate the impact of important events such as treatment outcomes. Thus, considering the intermittent course of depression over time, cycle lengths longer than 3 months may not be appropriate. The two studies we identified as using ISMs had cycle lengths of 1 and 2 months.

The authors of these two studies did not provide an explicit explanation regarding why they chose a microsimulation approach instead of a CMM. The ability of studies using ISM to model patient heterogeneity and memory can potentially improve the face validity of the model. The authors did not provide further information on other formal methods they used to validate the models.

3.3.4 Discrete-Event Simulation

DES modelling scored positively in seven of the 11 criteria (Table 4). It has been used extensively for other healthcare applications, such as surgical disciplines, but we identified only three studies of depression (7%). Using DES results in a microsimulation model in which each patient is simulated separately, similar to using an ISM. DES models do not use fixed cycles, instead, a constantly running simulation clock tracks time [47]. Thus, the model advances from one event to another and these interim periods (e.g. remission from depression) can be captured by referring to the simulation clock [47]. The time at which an event occurs is based on a survival function that models the time to that specific event.

Because of this flexible time management, an event can occur at any time, which is particularly useful when simulating recurrent disorders such as depression, where most patients are expected to experience multiple episodes. This flexibility in time can be approximated in state-transition models such as CMM and ISM. For instance, one of the included studies that used CMM had a cycle length of 2 weeks. Nevertheless, using very short cycle lengths may increase the running time of the state-transition models considerably.

Although patient interactions such as waiting lists, which are common in specialized depression care, can also

be implemented in the model, our literature search did not find any studies using this capability. The complexity of the model means the building and running time can be substantial. The expertise required to build and analyse DES models may be behind the limited use of this type of modelling in depression thus far. Nevertheless, dedicated software for DES models does exist, such as Simul8 [48], which one of the three studies used. The other two studies used Microsoft® Excel. In addition, data requirements are higher, posing a concern similar to that discussed for the ISM. However, DES is expected to represent the course of depression more naturally than DT or CMM because it is possible to model the effects of patient characteristics on the course of depression [49].

The authors of studies using DES models reported that they did so because of the complexity of the model they wanted to build and the ability to include patient heterogeneity and history [28, 29]. These advantages of DES modelling can enhance the face validity of the model. One [29] of the three studies described other methods of validation, including a workshop developed to validate the conceptual and mathematical models, testing of the model code and results by a second analyst, and error checking by the project team.

4 Discussion

This study aimed to review the methodological characteristics of model-based studies examining the cost effectiveness of treatments for depression and to evaluate which of the modelling methods was more appropriate for simulating the course of the disorder. The 41 studies had diverse methodological aspects and therefore it was difficult to compare them. Simpler models (i.e. DT and CMM) were used more frequently. However, microsimulation models (i.e. ISM and DES) appeared more appropriate when simulating the course of depression because they can track the personal history of each individual.

About half of the identified studies investigated antidepressants compared with a comparator treatment. However, practice guidelines recommend both psychological and pharmacological treatments for depressive disorders [50–52]. In addition, preliminary evidence from a meta-analysis showed that cognitive behavioural therapy and continuation of pharmacotherapy had similar long-term effects on depression [53]. Therefore, model-based economic evaluations of psychological treatments are necessary to examine the cost effectiveness of these interventions in the long term.

The included studies used various data sources to populate the models. Data from RCTs were commonly used to derive effect estimates. RCTs are methodologically sound but they sometimes have low external validity because of strict inclusion criteria [18, 54]. Meta-analysis of RCTs has increased precision and external validity than a single clinical trial, but the results may still be subject to publication bias [54, 55]. Observational studies typically have larger samples and higher external validity than RCTs [21]; however, the quality of the data is usually lower than in an RCT because it is difficult to control for confounding bias. Expert opinion is a controversial choice that may be used only when no other sources are available [18, 56]. We found 15 studies (37%) used expert opinion as the data source for clinical or cost inputs, which may be an indication of a lack of valid data for some parameters. The issue of the limited number of papers available in the literature presenting utility weights for different health states of depression has been addressed previously [57]. A recent systematic review and meta-analysis can provide a useful resource for decision analysts [58]. Overall, researchers should be cautious when using utility weights from the literature and ensure the weights they use are applicable to the specific population and health states they are evaluating.

The structure of the model must address the decision problem adequately [59]. The majority of the studies used relatively simple model structures with only a few health states, sometimes as few as two, with an average of four. Since depression is recurrent, with increased risk for suicide, it seems reasonable that the structure of the models needs to be relatively complex. Some key clinical events should be remission (ideally defined by a structured clinical interview) and a measure of symptom improvements (such as response or reliable change) and deterioration. Deterioration of depressive symptoms is often only partly included in the models as relapsing to depression after being in remission. However, it is possible that patients get worse while undergoing treatment for depression [60]. Moreover, depressive symptom severity is related to different utility weights, with more depressed patients having more deteriorated QoL [58]. Thus, accounting for severity of depression would improve the clinical representation of the models. Nevertheless, only three of the included studies used different states for different levels of depression severity. Overall, omitting important health states from the model structure may lead to overestimation or underestimation of the cost effectiveness of the treatment under study.

We searched the included papers for information on model validation. Model validation is crucial since it provides confidence to decision makers that the model accurately predicts the outcomes of interest [61]. We identified minimal information on validation methods. Although the body of literature on the importance of rigorous validation of statistical modelling is substantial [62], it appears not to

be common practice in the health economic modelling of depression treatments. This is a shortcoming that may cause flawed decision making.

We identified a larger number of more complex models in studies published after 2010. Afzali et al. [21] found one study using ISM, which was conducted in 2006, and no DES models. After 2010, another study that used ISM and three studies that used DES models were identified. In addition, we found three studies that used 'tunnel' states to model time in the CMM (conducted in 2010, 2013 and 2015). This may be an indication of a trend towards more complex modelling methods to investigate the cost effectiveness of depression treatments. More complex modelling methods may improve the face validity of the models.

Since depression is a chronic and recurrent disorder, the history of each patient is crucial for prognosis (e.g. more previous depressive episodes, more severe depression and comorbidity are all related to worse prognosis) [40]. Studies that used DES and ISM included patient heterogeneity (e.g. age, sex) and tracking of individual history (e.g. additional depressive episodes during the simulation) in the model estimations. This may improve not only the face validity of the models but also the validity of model predictions.

Nevertheless, we could not judge the validity of model prediction because it was not reported in the papers. On the contrary, studies that used DTs and CMM were not designed to include patient heterogeneity and history. Therefore, we consider DES and ISM the most appropriate methods to simulate the course of depression.

In addition, DES may have some advantages over ISM, such as increased flexibility in time management (no cycles are used since time runs constantly) and the ability to model patient interactions (e.g. waiting lists for access to treatment). However, our search did not detect any study using the ability of a DES model to simulate patient interactions. Only empirical comparisons between the different models will allow us to draw firm conclusions on whether important aspects of depression, such as patient heterogeneity and time-varying events, can be validly incorporated in cohort-based models without resulting in an unwieldy number of health states; this may be easier to incorporate in microsimulation models.

To our knowledge, such comparisons have not yet been conducted for depression treatments. One study has compared DES and CMM for HIV treatments [63] and one has compared them for early breast cancer treatments [64]. Simpson et al. [63] indicated that DES showed better predictive validity over 5 years. Karnon [64] concluded that DES represented the data more flexibly than did CMM but that this advantage was outweighed by the increased efficiency of CMM (i.e. short computational times, easier

to develop and analyse). Nevertheless, the extent to which these inferences can be generalized to the evaluation of depression treatments remains unknown.

In a recent review, Karnon and Afzali [65] presented an overview of the costs and benefits of using DES to model healthcare decision problems in general. The authors proposed four factors that decision analysts and researchers can use to evaluate whether using a DES model would be beneficial: baseline heterogeneity, disease progression as a continuous process, time-varying event rates, and prior events affect subsequent event rates [65]. When considering depression, it seems that all these factors apply. Initially, the depressed population is rather heterogeneous. Furthermore, depressive symptom severity is continuous, although sometimes it is conventionally modelled as discrete categories (i.e. mild, moderate, severe). In addition, time is crucial when modelling depression. The longer patients have depression the less likely they are to experience remission. Finally, the number of prior depressive episodes affects the probability of having new episodes. Thus, a DES model seems to be more suitable for modelling the course of depression than other more simple methods. However, the results of our review show that DES modelling is not yet commonly used for depression. This may limit the validity of the findings of existing studies.

More guidelines are available to help researchers and decision makers decide on the most suitable modelling method to address the problem of interest. The International Society for Pharmacoeconomics and Outcomes Research (ISPOR) Modeling Good Research Practices task force provided a thorough description of the advantages and disadvantages of the available modelling methods in a series of reports [22, 38, 47]. Finally, the ISPOR Dynamic Simulation Modeling Emerging Good Practices task force provided the SIMULATE checklist, which is a tool to assist modellers decide whether a dynamic modelling method, such as DES, is necessary or relatively simpler simulation methods, such as Markov models, are adequate to address the specific decision problem [66, 67]. In general, these guidelines are in accordance with our conclusions, indicating that chronic and recurrent health problems require more complex modelling techniques.

5 Conclusion

The present systematic review described the studies and compared the modelling techniques that have been used in the economic evaluation of depression treatments. There were substantial methodological differences between the studies, which decreased the comparability of the results. Since patient heterogeneity and the individual history of

each patient are important for the prognosis of depression, DES and ISM simulation methods may be more appropriate for a pragmatic representation of the course of depression. However, direct comparisons between the available modelling techniques are necessary to yield firm conclusions.

Compliance with Ethical Standards

Data availability statement The data used in this study are available as electronic supplementary material.

References

1. Mathers CD, Loncar D. Projections of global mortality and burden of disease from 2002 to 2030. PLoS Med. 2006;3(11):e442. doi:10.1371/journal.pmed.0030442.

2. Saarni SI, Suvisaari J, Sintonen H, Pirkola S, Koskinen S, Aromaa A, et al. Impact of psychiatric disorders on health-related quality of life: general population survey. Br J Psychiatry. 2007;190:326–32. doi:10.1192/bjp.bp.106.025106.

3. Papakostas GI, Petersen T, Mahal Y, Mischoulon D, Nierenberg AA, Fava M. Quality of life assessments in major depressive disorder: a review of the literature. Gen Hosp Psychiatry. 2004;26(1):13–7.

4. Stewart AL, Greenfield S, Hays RD, Wells K, Rogers WH, Berry SD, et al. Functional status and well-being of patients with chronic conditions. Results from the medical outcomes study. JAMA. 1989;262(7):907–13.

5. Kessler RC. The costs of depression. Psychiatr Clin North Am. 2012;35(1):1–14. doi:10.1016/j.psc.2011.11.005.

6. Greenberg PE, Fournier AA, Sisitsky T, Pike CT, Kessler RC. The economic burden of adults with major depressive disorder in the United States (2005 and 2010). J Clin Psychiatry. 2015;76(2):155–62. doi:10.4088/JCP.14m09298.

7. Wittchen HU, Jacobi F, Rehm J, Gustavsson A, Svensson M, Jonsson B, et al. The size and burden of mental disorders and other disorders of the brain in Europe 2010. Eur Neuropsychopharmacol. 2011;21(9):655–79. doi:10.1016/j.euroneuro.2011.07.018.

8. Olesen J, Gustavsson A, Svensson M, Wittchen HU, Jonsson B. The economic cost of brain disorders in Europe. Eur J Neurol. 2012;19(1):155–62. doi:10.1111/j.1468-1331.2011.03590.x.

9. Lépine J-P, Briley M. The increasing burden of depression. Neuropsychiatr Dis Treat. 2011;7(Suppl 1):3–7. doi:10.2147/NDT.S19617.

10. Goldberg D. The heterogeneity of "major depression". World Psychiatry. 2011;10(3):226–8.

11. Lieblich SM, Castle DJ, Pantelis C, Hopwood M, Young AH, Everall IP. High heterogeneity and low reliability in the diagnosis of major depression will impair the development of new drugs. Br J Psychiatr Open. 2015;1(2):e5–7. doi:10.1192/bjpo.bp.115.000786.

12. Simon GE, Perlis RH. Personalized medicine for depression: can we match patients with treatments? Am J Psychiatry. 2010;167(12):1445–55. doi:10.1176/appi.ajp.2010.09111680.

13. Sugarman MA. Are antidepressants and psychotherapy equally effective in treating depression? A critical commentary. J Mental Health (Abingd, Engl). 2016;25(6):475–478. doi:10.3109/09638237.2016.1139071.

14. Kessler RC, Berglund P, Demler O, Jin R, Koretz D, Merikangas KR, et al. The epidemiology of major depressive disorder: results from the National Comorbidity Survey Replication (NCS-R). JAMA. 2003;289(23):3095–105. doi:10.1001/jama.289.23.3095.

15. Solomon DA, Keller MB, Leon AC, Mueller TI, Lavori PW, Shea MT, et al. Multiple recurrences of major depressive disorder. Am J Psychiatry. 2000;157(2):229–33. doi:10.1176/appi.ajp.157.2.229.

16. Hasin DS, Goodwin RD, Stinson FS, Grant BF. Epidemiology of major depressive disorder: results from the National Epidemiologic Survey on Alcoholism and Related Conditions. Arch Gen Psychiatry. 2005;62(10):1097–106. doi:10.1001/archpsyc.62.10.1097.

17. Weinstein MC. Recent developments in decision-analytic modelling for economic evaluation. Pharmacoeconomics. 2006;24(11):1043–53.

18. Petrou S, Gray A. Economic evaluation using decision analytical modelling: design, conduct, analysis, and reporting. BMJ. 2011;342:d1766. doi:10.1136/bmj.d1766.

19. Le Lay A, Despiegel N, Francois C, Duru G. Can discrete event simulation be of use in modelling major depression? Cost Eff Resour Alloc. 2006;4:19. doi:10.1186/1478-7547-4-19.

20. Standfield L, Comans T, Scuffham P. Markov modeling and discrete event simulation in health care: a systematic comparison. Int J Technol Assess Health Care. 2014;30(2):165–72. doi:10.1017/s0266462314000117.

21. Afzali HHA, Karnon J, Gray J. A critical review of model-based economic studies of depression: modelling techniques, model structure and data sources. Pharmacoeconomics. 2012;30(6):461–82.

22. Caro JJ, Briggs AH, Siebert U, Kuntz KM. Modeling good research practices—overview: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-1. Value Health. 2012;15(6):796–803. doi:10.1016/j.jval.2012.06.012.

23. Zimovetz EA, Wolowacz SE, Classi PM, Birt J. Methodologies used in cost-effectiveness models for evaluating treatments in major depressive disorder: a systematic review. Cost Eff Resour Alloc. 2012;10(1):1.

24. IsHak WW, Greenberg JM, Balayan K, Kapitanski N, Jeffrey J, Fathy H, et al. Quality of life: the ultimate outcome measure of interventions in major depressive disorder. Harv Rev Psychiatry. 2011;19(5):229–39. doi:10.3109/10673229.2011.614099.

25. Brennan A, Chick SE, Davies R. A taxonomy of model structures for economic evaluation of health technologies. Health Econ. 2006;15(12):1295–310. doi:10.1002/hec.1148.

26. Heeg BM, Damen J, Buskens E, Caleo S, de Charro F, van Hout BA. Modelling approaches: the case of schizophrenia. Pharmacoeconomics. 2008;26(8):633–48.

27. Prukkanone B, Vos T, Bertram M, Lim S. Cost-effectiveness analysis for antidepressants and cognitive behavioral therapy for major depression in Thailand. Value Health. 2012;15(1 Suppl):S3–8. doi:10.1016/j.jval.2011.11.009.

28. Saylan M, Treur MJ, Postema R, Dilbaz N, Savas H, Heeg BM, et al. Cost-effectiveness analysis of aripiprazole augmentation treatment of patients with major depressive disorder compared to olanzapine and quetiapine augmentation in Turkey: a microsimulation approach. Value Health Reg Issues. 2013;2(2):171–80. doi:10.1016/j.vhri.2013.06.004.

29. Tosh J, Kearns B, Brennan A, Parry G, Ricketts T, Saxon D, et al. Innovation in health economic modelling of service improvements for longer-term depression: demonstration in a local health community. BMC Health Serv Res. 2013;13:150. doi:10.1186/1472-6963-13-150.

30. Dias S, Welton NJ, Sutton AJ, Ades AE. Evidence synthesis for decision making 5: the baseline natural history model. Med Decis Mak. 2013;33(5):657–70. doi:10.1177/0272989x13485155.

31. Golder S, Glanville J, Ginnelly L. Populating decision-analytic models: the feasibility and efficiency of database searching for individual parameters. Int J Technol Assess Health Care. 2005;21(3):305–11.

32. Briggs AH, Weinstein MC, Fenwick EA, Karnon J, Sculpher MJ, Paltiel AD. Model parameter estimation and uncertainty analysis: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force Working Group-6. Med Decis Making. 2012;32(5):722–32. doi:10.1177/0272989x12458348.

33. TreeAge Pro 2015. TreeAge Software, Williamstown, MA. https://www.treeage.com.

34. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS—a Bayesian modelling framework: concepts, structure, and extensibility. Stat Comput. 2000;10:325–37.

35. Mencacci C, Aguglia E, Biggio G, Cappellari L, Di Sciascio G, Fagiolini A, et al. C-QUALITY: cost and quality-of-life pharmacoeconomic analysis of antidepressants in major depressive disorder in Italy. Adv Ther. 2013;30(7):697–712. doi:10.1007/s12325-013-0046-z.

36. Nuijten MJ, Brignone M, Marteau F, den Boer JA, Hoencamp E. Cost-effectiveness of escitalopram in major depressive disorder in the Dutch health care setting. Clin Ther. 2012;34(6):1364–78. doi:10.1016/j.clinthera.2012.04.006.

37. Trivedi MH, Wan GJ, Mallick R, Chen J, Casciano R, Geissler EC, et al. Cost and effectiveness of venlafaxine extended-release and selective serotonin reuptake inhibitors in the acute phase of outpatient treatment for major depressive disorder. J Clin Psychopharmacol. 2004;24(5):497–506. doi:10.1097/01.jcp.0000138769.61600.e4.

38. Siebert U, Alagoz O, Bayoumi AM, Jahn B, Owens DK, Cohen DJ, et al. State-transition modeling: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-3. Value Health. 2012;15(6):812–20. doi:10.1016/j.jval.2012.06.014.

39. Moradveisi L, Huibers M, Renner F, Arntz A. The influence of patients' preference/attitude towards psychotherapy and antidepressant medication on the treatment of major depressive disorder. J Behav Ther Exp Psychiatry. 2014;45(1):170–7. doi:10.1016/j.jbtep.2013.10.003.

40. Hardeveld F, Spijker J, De Graaf R, Nolen WA, Beekman AT. Recurrence of major depressive disorder and its predictors in the general population: results from the Netherlands Mental Health Survey and Incidence Study (NEMESIS). Psychol Med. 2013;43(1):39–48. doi:10.1017/s0033291712002395.

41. Keller MB, Boland RJ. Implications of failing to achieve successful long-term maintenance treatment of recurrent unipolar major depression. Biol Psychiatry. 1998;44(5):348–60.

42. Benedict Á, Arellano J, De Cock E, Baird J. Economic evaluation of duloxetine versus serotonin selective reuptake inhibitors and venlafaxine XR in treating major depressive disorder in Scotland. J Affect Disord. 2010;120(1–3):94–104. doi:10.1016/j.jad.2009.04.017.

43. Olgiati P, Bajo E, Bigelli M, De Ronchi D, Serretti A. Should pharmacogenetics be incorporated in major depression treatment? Economic evaluation in high- and middle-income European countries. Prog Neuropsychopharmacol Biol Psychiatry. 2012;36(1):147–54. doi:10.1016/j.pnpbp.2011.08.013.

44. Siskind D, Araya R, Kim J. Cost-effectiveness of improved primary care treatment of depression in women in Chile. Br J Psychiatry. 2010;197(4):291–6. doi:10.1192/bjp.bp.109.068957.

45. Wang PS, Patrick A, Avorn J, Azocar F, Ludman E, McCulloch J, et al. The costs and benefits of enhanced depression care to employers. Arch Gen Psychiatry. 2006;63(12):1345–53. doi:10.1001/archpsyc.63.12.1345.

46. Sonnenberg FA, Beck JR. Markov models in medical decision making: a practical guide. Med Decis Mak. 1993;13(4):322–38.

47. Karnon J, Stahl J, Brennan A, Caro JJ, Mar J, Moller J. Modeling using discrete event simulation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-4. Value Health. 2012;15(6):821–7. doi:10.1016/j.jval.2012.04.013.

48. Elder M. DES view on simulation modelling: SIMUL8. In: Brailsford S, Churilov L, Dangerfield B, editors. Discrete-event simulation and system dynamics for management decision making. Chichester: Wiley; 2014. p. 199–214.

49. Caro JJ. Pharmacoeconomic analyses using discrete event simulation. Pharmacoeconomics. 2005;23(4):323–32.

50. American Psychiatric Association. American Psychiatric Association Practice Guidelines for the treatment of psychiatric disorders: compendium 2006. Arlington: American Psychiatric Publishing; 2006.

51. National Collaborating Centre for Mental Health (UK). Depression: the treatment and management of depression in adults (updated edition). NICE Clinical Guidelines, no. 90. Leicester: British Psychological Society; 2010.

52. Ellis P. Australian and New Zealand clinical practice guidelines for the treatment of depression. Aust N Z J Psychiatry. 2004;38(6):389–407. doi:10.1111/j.1440-1614.2004.01377.x.

53. Cuijpers P, Hollon SD, van Straten A, Bockting C, Berking M, Andersson G. Does cognitive behaviour therapy have an enduring effect that is superior to keeping patients on continuation pharmacotherapy? A meta-analysis. BMJ Open. 2013;3:e002542. doi:10.1136/bmjopen-2012-002542.

54. Nuijten MJ. The selection of data sources for use in modelling studies. Pharmacoeconomics. 1998;13(3):305–16.

55. Cuijpers P, van Straten A, Bohlmeijer E, Hollon SD, Andersson G. The effects of psychotherapy for adult depression are overestimated: a meta-analysis of study quality and effect size. Psychol Med. 2010;40(2):211–23. doi:10.1017/s0033291709006114.

56. Iqbal SU, Prashker M. Pharmacoeconomic evaluation of antidepressants: a critical appraisal of methods. Pharmacoeconomics. 2005;23(6):595–606.

57. Sonntag M, Konig HH, Konnopka A. The estimation of utility weights in cost-utility analysis for mental disorders: a systematic review. Pharmacoeconomics. 2013;31(12):1131–54. doi:10.1007/s40273-013-0107-9.

58. Mohiuddin S, Payne K. Utility values for adults with unipolar depression: systematic review and meta-analysis. Med Decis Mak. 2014;34(5):666–85.

59. Roberts M, Russell LB, Paltiel AD, Chambers M, McEwan P, Krahn M. Conceptualizing a model: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-2. Med Decis Mak. 2012;32(5):678–89. doi:10.1177/0272989x12454941.

60. Ebert DD, Donkin L, Andersson G, Andrews G, Berger T, Carlbring P, et al. Does Internet-based guided-self-help for depression cause harm? An individual participant data meta-analysis on deterioration rates and its moderators in randomized controlled trials. Psychol Med. 2016;46(13):2679–93. doi:10.1017/s0033291716001562.

61. Eddy DM, Hollingworth W, Caro JJ, Tsevat J, McDonald KM, Wong JB. Model transparency and validation: a report of the ISPOR-SMDM Modeling Good Research Practices Task Force-7. Med Decis Mak. 2012;32(5):733–43. doi:10.1177/0272989x12454579.

62. Steyerberg EW, Vergouwe Y. Towards better clinical prediction models: seven steps for development and an ABCD for validation. Eur Heart J. 2014;35(29):1925–31. doi:10.1093/eurheartj/ehu207.

63. Simpson KN, Strassburger A, Jones WJ, Dietz B, Rajagopalan R. Comparison of Markov model and discrete-event simulation techniques for HIV. Pharmacoeconomics. 2009;27(2):159–65. doi:10.2165/00019053-200927020-00006.

64. Karnon J. Alternative decision modelling techniques for the evaluation of health care technologies: Markov processes versus discrete event simulation. Health Econ. 2003;12(10):837–48. doi:10.1002/hec.770.

65. Karnon J, Haji Ali Afzali H. When to use discrete event simulation (DES) for the economic evaluation of health technologies? A review and critique of the costs and benefits of DES. Pharmacoeconomics. 2014;32(6):547–58. doi:10.1007/s40273-014-0147-9.

66. Marshall DA, Burgos-Liz L, Ijzerman MJ, Crown W, Padula WV, Wong PK, et al. Selecting a dynamic simulation modeling method for health care delivery research-part 2: report of the ISPOR Dynamic Simulation Modeling Emerging Good Practices Task Force. Value Health. 2015;18(2):147–60. doi:10.1016/j.jval.2015.01.006.

67. Marshall DA, Burgos-Liz L, IJzerman MJ, Osgood ND, Padula WV, Higashi MK, et al. Applying dynamic simulation modeling methods in health care delivery research-the SIMULATE checklist: report of the ISPOR simulation modeling emerging good practices task force. Value Health. 2015;18(1):5–16. doi:10.1016/j.jval.2014.12.001.

68. Armstrong EP, Skrepnek GH, Haim Erder M. Cost-utility comparison of escitalopram and sertraline in the treatment of major depressive disorder. Curr Med Res Opin. 2007;23(2):251–8. doi:10.1185/030079907x159498.

69. Armstrong EP, Malone DC, Erder MH. A Markov cost-utility analysis of escitalopram and duloxetine for the treatment of major depressive disorder. Curr Med Res Opin. 2008;24(4):1115–21. doi:10.1185/030079908x273309.

70. Aziz M, Mehringer AM, Mozurkewich E, Razik GN. Cost-utility of 2 maintenance treatments for older adults with depression who responded to a course of electroconvulsive therapy: results from a decision analytic model. Can J Psychiatry. 2005;50(7):389–97.

71. Beil H, Beeber LS, Schwartz TA, Lewis G. Cost-effectiveness of alternative treatments for depression in low-income women. J Mental Health Policy Econ. 2013;16(2):55–65.

72. Evans-Lacko S, Koeser L, Knapp M, Longhitano C, Zohar J, Kuhn K. Evaluating the economic impact of screening and treatment for depression in the workplace. Eur Neuropsychopharmacol. 2016;26(6):1004–13. doi:10.1016/j.euroneuro.2016.03.005.

73. Francois CSH, Toumi M. Introduction of escitalopram, a new SSRI in Finland: comparison of cost-effectiveness between the other SSRIs and SNRI for the treatment of depression and estimation of the budgetary impact. J Med Econ. 2002;5:91–107.

74. Greenhalgh J, Knight C, Hind D, Beverley C, Walters S. Clinical and cost-effectiveness of electroconvulsive therapy for depressive illness, schizophrenia, catatonia and mania: systematic reviews and economic modelling studies. Health Technol Assess. 2005;9(9):1–156, iii–iv.

75. Kaltenthaler E, Brazier J, De Nigris E, Tumur I, Ferriter M, Beverley C et al. Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. Health Technol Assess. 2006;10(33):iii, xi–xiv, 1–168.

76. Khoo AL, Zhou HJ, Teng M, Lin L, Zhao YJ, Soh LB, et al. Network Meta-Analysis and Cost-Effectiveness Analysis of New Generation Antidepressants. CNS Drugs. 2015;29(8):695–712. doi:10.1007/s40263-015-0267-6.

77. Koeser L, Dobbin A, Ross S, McCrone P. Economic evaluation of audio based resilience training for depression in primary care. J Affect Disord. 2013;149(1–3):307–12. doi:10.1016/j.jad.2013.01.044.

78. Koeser L, Donisi V, Goldberg DP, McCrone P. Modelling the cost-effectiveness of pharmacotherapy compared with cognitive–behavioural therapy and combination therapy for the treatment of moderate to severe depression in the UK. Psychol Med. 2015;45(14):3019–31. doi:10.1017/S0033291715000951.

79. Kozel FA, George MS, Simpson KN. Decision analysis of the cost-effectiveness of repetitive transcranial magnetic stimulation versus electroconvulsive therapy for treatment of nonpsychotic severe depression. CNS Spectr. 2004;9(6):476–82.

80. Leelahanaj T. The cost-effectiveness of aripiprazole as adjunctive therapy in major depressive disorder: Thai economic model. J Med Assoc Thail. 2010;93(Suppl 6):S43–50.

81. Leelahanaj T. Developing thai economic model to study cost-effectiveness of switching to bupropion compared to combination with bupropion after the failure of an SSRI for major depressive disorder. J Med Assoc Thai. 2010;93(Suppl 6):S35–42.

82. Lenox-Smith A, Greenstreet L, Burslem K, Knight C. Cost effectiveness of venlafaxine compared with generic fluoxetine or generic amitriptyline in major depressive disorder in the UK. Clin Drug Investig. 2009;29(3):173–84. doi:10.2165/00044011-200929030-00004.

83. Maniadakis N, Kourlaba G, Mougiakos T, Chatzimanolis I, Jonsson L. Economic evaluation of agomelatine relative to other antidepressants for treatment of major depressive disorders in Greece. BMC Health Serv Res. 2013;13:173. doi:10.1186/1472-6963-13-173.

84. Naversnik K, Mrhar A. Cost-effectiveness of a novel e-health depression service. Telemed J E Health. 2013;19(2):110–6. doi:10.1089/tmj.2012.0081.

85. Nguyen KH, Gordon LG. Cost-effectiveness of repetitive transcranial magnetic stimulation versus antidepressant therapy for treatment-resistant depression. Value Health. 2015;18(5):597–604. doi:10.1016/j.jval.2015.04.004.

86. Nordstrom G, Despiegel N, Marteau F, Danchenko N, Maman K. Cost effectiveness of escitalopram versus SNRIs in second-step treatment of major depressive disorder in Sweden. J Med Econ. 2010;13(3):516–26. doi:10.3111/13696998.2010.506371.

87. Olgiati P, Bajo E, Bigelli M, Montgomery S, Serretti A. Challenging sequential approach to treatment resistant depression: cost-utility analysis based on the Sequenced Treatment Alternatives to Relieve Depression (STAR()D) trial. Eur Neuropsychopharmacolean. 2013;23(12):1739–46. doi:10.1016/j.euroneuro.2013.08.008.

88. Ramsberg J, Asseburg C, Henriksson M. Effectiveness and cost-effectiveness of antidepressants in primary care: a multiple treatment comparison meta-analysis and cost-effectiveness model. PLoS One. 2012;7(8):e42003. doi:10.1371/journal.pone.0042003.

89. Sado M, Knapp M, Yamauchi K, Fujisawa D, So M, Nakagawa A, et al. Cost-effectiveness of combination therapy versus antidepressant therapy for management of depression in Japan. Aust N Z J Psychiatry. 2009;43(6):539–47. doi:10.1080/00048670902873664.

90. Serretti A, Olgiati P, Bajo E, Bigelli M, De Ronchi D. A model to incorporate genetic testing (5-HTTLPR) in pharmacological treatment of major depressive disorders. World J Biol Psychiatry. 2011;12(7):501–15. doi:10.3109/15622975.2011.572998.

91. Simon J, Pilling S, Burbeck R, Goldberg D. Treatment options in moderate and severe depression: decision analysis supporting a clinical guideline. Br J Psychiatry. 2006;189:494–501. doi:10.1192/bjp.bp.105.014571.

92. Siskind D, Baingana F, Kim J. Cost-effectiveness of group psychotherapy for depression in Uganda. J Ment Health Policy Econ. 2008;11(3):127–33.

93. Sobocki P, Ekman M, Agren H, Jonsson B, Rehnberg C. Model to assess the cost-effectiveness of new treatments for depression. Int J Technol Assess Health Care. 2006;22(4):469–77. doi:10.1017/s0266462306051397.

94. Sobocki P, Ekman M, Ovanfors A, Khandker R, Jonsson B. The cost-utility of maintenance treatment with venlafaxine in patients with recurrent major depressive disorder. Int J Clin Pract. 2008;62(4):623–32. doi:10.1111/j.1742-1241.2008.01711.x.

95. Solomon D, Adams J, Graves N. Economic evaluation of St. John's wort (Hypericum perforatum) for the treatment of mild to moderate depression. J Affect Disord. 2013;148(2–3):228–34.

96. Solomon D, Proudfoot J, Clarke J, Christensen H. e-CBT (myCompass), antidepressant medication, and face-to-face psychological treatment for depression in Australia: a cost-effectiveness comparison. J Med Internet Res. 2015;17(11):e255. doi:10.2196/jmir.4207.

97. Sullivan PW, Valuck R, Saseen J, MacFall HM. A comparison of the direct costs and cost effectiveness of serotonin reuptake inhibitors and associated adverse drug reactions. CNS Drugs. 2004;18(13):911–32.

98. Vallejo-Torres L, Castilla I, González N, Hunter R, Serrano-Pérez P, Perestelo-Pérez L. Cost-effectiveness of electroconvulsive therapy compared to repetitive transcranial magnetic stimulation for treatment-resistant severe depression: a decision model. Psychol Med. 2015;45(7):1459–70. doi:10.1017/S0033291714002554.