Journal of Affective Disorders 335 (2023) 313–321

Contents lists available at ScienceDirect

Journal of Affective Disorders

journal homepage: www.elsevier.com/locate/jad

Research paper

Cost-effectiveness of long-term psychoanalytic psychotherapy for treatment-resistant depression: RCT evidence from the Tavistock Adult Depression Study (TADS)

Leonardo Koeser [a], Felicitas Rost [b,c,*], Andrea Gabrio [d], Thomas Booker [c], David Taylor [c], Peter Fonagy [e], David Goldberg [f], Martin Knapp [g], Paul McCrone [a]

[a] King's Health Economics, Health Services and Population Research Department, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[b] The Open University, School of Psychology and Psychotherapy, Faculty of Arts and Social Sciences, Milton Keynes, UK
[c] Tavistock and Portman NHS Foundation Trust, London, UK
[d] Department of Methodology and Statistics, Faculty of Health, Medicine and Life Sciences, Maastricht University, Maastricht, the Netherlands
[e] Research Department of Clinical, Educational and Health Psychology, University College London, UK
[f] Institute of Psychiatry, Psychology and Neuroscience, King's College London, London, UK
[g] Personal Social Services Research Unit, London School of Economics and Political Science, London, UK

ARTICLE INFO

Keywords:
Randomized controlled trial
Treatment-resistant depression
Psychoanalytic psychotherapy
Cost-utility analysis
Economic evaluation

ABSTRACT

Background: Treatment-resistant depression (TRD) accounts for a large fraction of the burden of depression. The interventions currently used are mostly pharmacological and short-term psychotherapies, but their effectiveness is limited. The Tavistock Adult Depression Study found evidence for the effectiveness of long-term psychoanalytic psychotherapy (LTPP) plus treatment as usual (TAU), versus TAU alone, for TRD. Even after a 2-year follow-up, moderate effect sizes were sustained. This study assessed the cost-effectiveness of this LTPP + TAU.
Methods: We conducted a within-trial economic evaluation using a Bayesian framework.
Results: Quality-adjusted life years (QALYs) were 0.16 higher in the LTPP + TAU group compared with TAU. The direct cost of LTPP was £5500, with no substantial compensating savings elsewhere. Overall, average health and social care costs in the LTPP + TAU group were £5000 more than in the TAU group, employment rates were unchanged, and effects on other non-healthcare costs were uncertain. Accordingly, the incremental cost-effectiveness ratio was ≈£33,000/QALY; the probability that LTPP + TAU was cost-effective at a willingness to pay of £20,000/QALY was 18 %.
Limitations: The sample size of this study was relatively small, and the fraction of missing service-use data was approximately 50 % at all time points. The study was conducted at a single site, potentially reducing generalizability.
Conclusions: Although LTPP + TAU was found to be clinically effective for treating TRD, it was not found to be cost-effective compared with TAU. However, given the sustained effects over the follow-up period it is likely that the time horizon of this study was too short to capture all benefits of LTPP augmentation.

1. Introduction

Persisting forms of depression are often referred to as 'treatment-resistant depression' (TRD). Although estimates of prevalence vary according to diagnostic criteria and population, typically between 15 % and 30 % of people with a major depressive disorder (MDD) will not respond to the initial, and even multiple, treatments offered (Jobst et al., 2016; McIntyre et al., 2014; Rush et al., 2006; Thomas et al., 2013) As well as prolonged depressive symptoms, many individuals experience other mental and physical health morbidities, difficulties in interpersonal relationships, and poor occupational functioning (Jobst et al., 2016; Nemeroff et al., 2003; Satyanarayana et al., 2009). This often leads to particularly poor quality of life (Johnston et al., 2019; Mrazek et al., 2014). Compared with people who respond to treatment for their

* Corresponding author at: Tavistock and Portman NHS Foundation Trust, Portman Clinic, 120 Belsize Lane, London NW3 5BA, UK.
E-mail address: felicitas.rost@open.ac.uk (F. Rost).

https://doi.org/10.1016/j.jad.2023.04.109
Received 14 November 2022; Received in revised form 21 April 2023; Accepted 28 April 2023
Available online 8 May 2023

depression, patients with TRD have substantially higher healthcare costs and productivity losses (Johnston et al., 2019). Thus, if an intervention is effective, there is potential for its costs to be at least partially offset by savings made elsewhere.

To address the paucity of research guiding the clinical management of individuals with TRD, Fonagy et al. (2015) conducted the Tavistock Adult Depression Study (TADS), a randomized controlled trial (RCT) to examine the effectiveness of augmenting treatment as usual (TAU) for primary-care patients diagnosed with TRD with long-term psychoanalytic psychotherapy (LTPP). TAU consisted of a range of brief psychotherapies recommended by the UK National Institute for Health and Care Excellence guidelines that were current at the time (NICE, 2010) and LTPP consisted of 18 months of once-weekly psychoanalytic psychotherapy offered at the Tavistock Clinic, an outpatient National Health Service (NHS) clinic. The authors found statistically significant improvements associated with LTPP in terms of depression severity as measured by the Hamilton Rating Scale for Depression (HRSD-17) and the Beck Depression Inventory (BDI-II), and in terms of Diagnostic and Statistical Manual Version IV (DSM-IV) diagnostic criteria (APA, 1994; Beck et al., 1996; Hamilton, 1960). At the end of the 2-year follow-up period, 44 % of those receiving LTPP + TAU no longer met diagnostic criteria for MDD compared with 4 % in the group receiving TAU only. Similar statistically significant differences between the two groups were found with respect to subjective well-being, symptoms, functioning, and level of clinical risk as measured by the Clinical Outcomes Routine Evaluation – Outcome Measure (CORE-OM) (Evans, 2000). Overall, these findings have added to the growing evidence base of the effectiveness of LTPP for depression (Leichsenring et al., 2015). However, as far as we know, the cost-effectiveness of LTPP for treating TRD has not yet been evaluated. Available economic evaluations of LTPP are predominantly based on observational data and involve non-TRD populations or populations with less severe forms of TRD, as well as interventions and comparators that do not generalize well to clinical practice in the NHS. The only existing trial-based RCT (Maljanen et al., 2016), for example, included individuals with anxiety or less severe forms of depression and compared a more lengthy course of LTPP and more lengthy courses of psychotherapies in the comparator arms. Given the substantial clinical and economic impacts of TRD, and the reported benefits of augmenting TAU with LTPP in the TADS, an economic evaluation of LTPP is clearly indicated. Investment in other provisions might yield more, or less, cost-effective benefits. Therefore, the aim was to undertake an economic evaluation, whose method and findings we report here.

2. Method

2.1. Data and study design

We carried out a cost-utility analysis embedded within TADS, a pragmatic single-centre RCT evaluating 129 participants over a 3.5-year time period. In the base case analysis, we took a healthcare service perspective on costs, and a patient perspective on the benefits of treatment. All comparisons, both those in the parent RCT and in this economic analysis, follow an intention-to-treat (ITT) approach. Ethical approval was sought and approved by the Institutional Review Board of NHS West Midlands Research Ethics Committee (MREC02/07/035).

The TADS trial protocol is reported by Taylor et al. (2012) and its outcome findings are reported by Fonagy et al. (2015) and Rost et al. (2019). Participants, recruited from primary care in London, UK, met research diagnostic criteria for MDD and/or dysthymia, a minimum of a 2-year history of depression, and at least two previous failed treatment attempts (of which one must have been an antidepressant). Acknowledging different definitions of TRD (Fekadu et al., 2018), the definition utilized in the TADS met generally accepted clearly defined criteria. In respect of physical and mental health comorbidities, developmental trauma and personality disorder, participants were coincidentally found to meet the criteria for 'complex and persistent depression' as well (Jobst et al., 2016; Ruhé et al., 2012).

2.2. TADS interventions

Detailed inventories of the treatments administered are reported elsewhere (Fonagy et al., 2015; Rost et al., 2019). The TAU included a range of short-term treatments prescribed by the participant's primary care doctor, referred to in the UK as a general practitioner (GP). These treatments were those recommended by the 2010 UK national treatment guidelines for depression (NICE, 2010). Participants allocated to the LTPP + TAU group were offered 60 weekly individual sessions of psychoanalytic psychotherapy over an 18-month period, provided by the Tavistock and Portman NHS Foundation Trust, in addition to TAU. In both groups, a majority (70–80 %) of individuals received antidepressant medication both during the treatment and over the follow-up period.

2.3. Measures of effectiveness

In England and Wales, quality-adjusted life years (QALYs) are the preferred outcome measure for economic evaluations. They can be calculated by converting changes in the relevant outcome study's measures by using a previously established algorithm (Whitehead and Ali, 2010) Our study utilizes the TADS main study data of the CORE-OM (Evans, 2000) and the HRSD-17 (Hamilton, 1960) collected at baseline, at 3-month intervals during the treatment, at the end of treatment, and then at 24, 30, and 42 months after treatment. The HRSD-17 is a commonly used, observer-rated depression-rating scale (Williams, 2001). It was the TADS primary outcome measure. The ratings were done by two independent raters blind to participants' treatment allocation. The intraclass correlation coefficient of inter-rater reliability was 0.89. Regrettably, to date no algorithm converting HRSD-17 scores to QALYs exists. A measure has been developed that derives QALY utility scores using the CORE-6D subset of the CORE-OM – a 34-item self-report measure of well-being, functioning, risk, and problems/symptoms widely used in the UK to evaluate psychological therapy services (Mavranezouli et al., 2012). We therefore used the CORE-6D subset of the CORE-OM in this economic analysis to derive QALY utility scores. We also report the HRSD-17 estimates for the purpose of comparison. Since the effects of treatments for depression on employment rates are important to policy decisions (Clark, 2011), we estimated the impact of LTPP on participants' working status in a secondary analysis.

2.4. Recording service use and calculating costs

We obtained information on healthcare service use other than LTPP, receipt of unpaid care receipt, and lost days of work due to depression from two sources: firstly, from the self-report Client Service Receipt Inventory (CSRI), which was collected every 6 months after baseline until 30 months, and once more at 42 months (Beecham and Knapp, 2001); secondly, from the service use data held by the patients' GP records, for which separate consent was sought from the participants. To be consistent in the service use captured by the two sources, we started the analysis from the beginning of treatment and included only medical records data recorded with the time period specified in the respective CSRIs. We assumed under-reporting of service use to be more likely than over-reporting in our study protocol. Therefore, in the event of discrepancies between medical records and self-report data, we prespecified that the higher value would be used.

The cost of medications was not included: there were only small differences in antidepressant use between the two arms of the trial, and generally, the cost of medications for depression is low (Edwards et al., 2013; Maljanen et al., 2016).

The number of LTPP sessions attended by the patients in the LTPP + TAU group was recorded separately. We costed LTPP at £109 per

attended session and £546 for the three initial assessment interviews by using reimbursement data obtained from the care provider (at that time, the patients' primary care trust). We used standard Department of Health 2009–2010 reference costs to value hospital service use, and the unit cost compendium by Curtis (2010) to value community service use other than LTPP. We adjusted all costs to 2014 prices using the hospital and community service index; we discounted both costs and outcomes at a rate of 3.5 % (Curtis and Burns, 2018).

In our primary analysis, we adopted a healthcare service perspective. In a secondary analysis, we calculated the cost of unpaid care and days off work. For this, we adopted a human capital approach based on average wages (approximately £13/h) (Bovill, 2014).

2.5. Statistical and cost-effectiveness analyses

We adopted a Bayesian framework in favor of the more conventional frequentist approach. The reasons for doing so are twofold: Firstly, it allowed us to estimate the joint uncertainty of five models, namely a model for costs, CORE-6D, HRSD-17, survival and treatment costs in a coherent, unified framework (Baio, 2012) which would be much more challenging to achieve using a frequentist approach. Secondly, it permitted us to handle missing data in a way that allowed the features of some of the missing variables while being congruent with the analysis model (Lambert et al., 2008). Again, using a frequentist framework, it would be difficult To multiply impute missing data as it has both multilevel structure and is assumed to have a skewed (e.g., gamma) distribution while providing stable estimates given the relatively small number of non-missing observations at each time point. Supplementary material 2 reports the technical details of the statistical approach we used. In summary, in our base case analysis, we used a longitudinal model with vague priors. To model CORE-6D utilities and HRSD-17 we used normal distributions, and a gamma distribution to model costs. We undertook baseline adjustment and assumed that data were missing at random. The syntax used to implement this model in JAGS and R (Plummer, 2016) is included in the Supplementary material.

We estimated the incremental cost-effectiveness ratio (ICER) between the two treatments by dividing the difference in mean costs by the difference in mean QALYs. We displayed the joint uncertainty surrounding these estimates using credible ellipses on a cost-effectiveness plane. The vertical axis of cost-effectiveness planes shows differences in costs whereas the horizontal axis shows differences in effects between the two treatment arms relative to current practice. Depending on the sign of the differences in costs and effects one typically distinguishes between four quadrants. If QALYs are the measure of effect, results in each quadrant can be interpreted as follows:

- North-east (NE) quadrant: The intervention is more costly and more effective than current practice.
- South-east (SE) quadrant: The intervention is less costly and more effective than current practice.
- South-west (SW) quadrant: The intervention is less costly and less effective than current practice.
- North-east (NE) quadrant: The intervention is more costly and less effective than current practice.

Interventions that are likely to fall within the SE quadrant should be adopted because they are superior in all aspects whereas those falling within the NE quadrant should not be adopted because they yield no added value. Those falling within the NE or SW quadrants will require a weighing of costs relative to benefits. When HRSD-17 is used as an outcome on cost-effectiveness planes, the interpretation of eastern and western quadrants is reversed because low rather than high HRSD-17 scores are desirable. We produced cost-effectiveness acceptability curves to show the probability that the intervention is cost-effective at different levels of willingness to pay per QALY improvement. This indicates the decision uncertainty facing the decision maker (Fenwick et al., 2004).

2.6. Sensitivity analyses

Pre-specified in our analysis plan, we explored:

I. The effects of altering the cost of the intervention by ±25 %
II. Using minimum rather than average wages to cost lost employment and receipt of unpaid care (approximately £6/h in 2010) (DBIS, 2011).

In addition, we explored:

III. Alternative approaches to costing service use, including:
   a. Determining which of the two cost sources was higher at the subcategory level (e.g., hospital costs)
   b. Aggregating these subgroup maxima rather than first aggregating the subcategories and then determining which of the two cost sources was higher, as in the base case analysis
   c. Calculating costs based only on data from the CSRI
   d. Calculating costs based only on service use data from medical records
IV. The effect of reducing the CORE-6D score below the predicted value for all time points that had missing data. This was to assess the robustness of the results in relation to data 'missing not at random' (MNAR), which may have occurred because depressed patients may miss appointments or drop out of treatment intermittently as a result of feeling unwell. We could not find evidence to inform the size of the adjustment necessary due to MNAR data. We chose to decrease utility values by 0.1 below the value predicted by the model in both trial conditions
V. The impacts of using a log-normal rather than a gamma distribution to model cost data.

3. Results

3.1. Characteristics of the sample

The ITT sample of the TADS consisted of 129 patients whose mean age was 44 years; 67 were randomized to LTPP + TAU and 62 to TAU. 85 (66 %) were women. As well as the minimum 2-year history of MDD and at least two previous failed treatment attempts, 81 % met the criteria for early-onset dysthymia. The average lifetime duration of depression was 25 years (standard deviation (SD) = 12); the average length of the current MDD episode was 3.7 years (SD = 3). Of the participants, 70 (54 %) were unemployed. The mean HRSD-17 score at baseline was 20, and the mean CORE-6D utility score was 0.54. Supplementary Figs. 1 to 8 show the missing data patterns graphically. Overall, approximately 35 % of CORE-OM, 27 % of HRSD-17, 50 % of self-report service use, and 17 % of medical records data were missing; there were no pronounced differences in missingness rates between groups.

3.2. Resource use and employment rates

Fig. 1(a) shows the contacts with mental healthcare professionals by study participants. Those in the LTPP + TAU group attended on average 48 sessions of psychoanalytic psychotherapy, as well as three assessment sessions and three review sessions. In addition, they received on average four sessions of individual non-psychoanalytic psychotherapy as prescribed by their GP. Those in the TAU group had on average 12 sessions of non-psychoanalytic individual psychotherapy, and one person's allocation to the TAU group was countermanded for clinical reasons. For

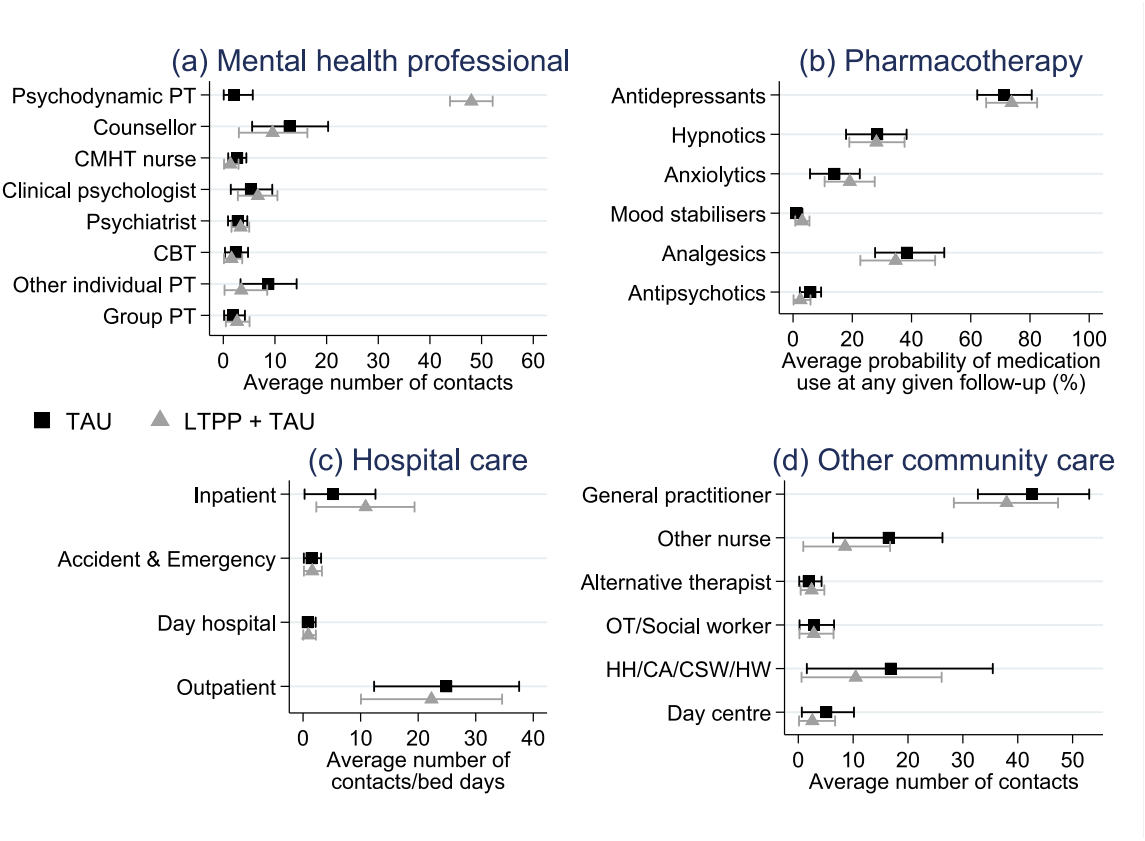

Fig. 1. Selected healthcare received over 3.5-year follow-up.
TAU: Treatment as usual; LTPP: Long-term psychoanalytic psychotherapy; PT: psychotherapist; CBT: cognitive-behavioural therapist; CMHT: community mental health team; OT: Occupational therapist; HH: home help; CA: care assistant; CSW: community support worker; HW: home worker.

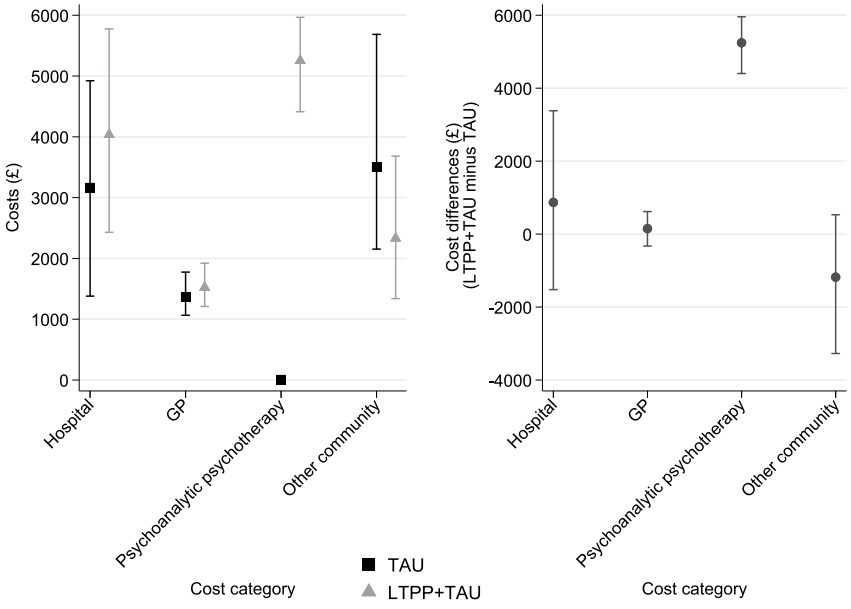

Fig. 2. Healthcare costs over the 3.5-year follow-up.
TAU: treatment as usual; LTPP: long-term psychoanalytic psychotherapy.

both the LTPP + TAU and TAU groups, contacts with counsellors and clinical psychologists averaged 10 sessions; those with psychiatrists, community mental health team nurses, and group psychotherapy sessions averaged less than five.

Fig. 1(b) shows the use of medication among participants in the two groups. Medication use was similar in the two groups. Antidepressants were the most commonly prescribed medication, with about 70 % of the patients receiving them at each measurement stage.

Fig. 1(c) shows the amount of hospital care received by the two groups. As shown, there was little difference between the groups. However, the average number of inpatient days was higher in the LTPP + TAU group than in the TAU group (11 vs. 3 bed days). The credible interval (CrI) around this estimate was wide, which indicates that the difference was very uncertain.

Fig. 1(d) shows contacts with other community services, including with GPs, non-CMHT nurses, day-care centres, and home helpers. We estimated that those in the TAU group had more contacts than those in the LTPP + TAU group. However, CrIs were wide in all cases, again questioning the certainty of the differences.

In summary, those in the LTPP + TAU group reported receiving informal care for 2.7 h per week on average, compared with 3.4 h reported by those in TAU (mean difference 1, 95 % CrI −2 to 4). With respect to employment status, there was little change across both groups over the course of the trial. We estimated that more individuals in the LTPP arm were in employment compared with those receiving TAU (52 % vs. 43 %, mean difference −9 %, 95 % CrI −19 to 2). On average, nine working days were reported to have been lost due to illness over the trial period in both trial groups (mean difference 0, 95 % CrI −13 to 14).

3.3. Costs

Fig. 2 shows the average cost of LTPP. We estimated this to be £5468 (95 % CrI £4854 to £6087). On average, the cost of hospital care in the LTPP + TAU group was £2800 higher than that in the TAU group, although this estimate was subject to considerable uncertainty (95 % CrI −£1176 to £9265). GP costs were almost identical in the two groups despite the TAU group's greater number of contacts (mean difference £26, 95 % CrI −£810 to £1024). This is explained by differences in the length and type of the contacts. The cost of other community care was substantially lower in the LTPP + TAU group. This estimate was associated with a large degree of uncertainty (mean difference –£1565, 95 % CrI −£4880 to £881).

Fig. 3 shows the trajectory of mean healthcare costs, not including LTPP, over the course of the study. No clear temporal patterns were found in the differences between the two groups. The cost difference between the groups was due mainly to differences in the cost of treatment. In the base case analysis, we estimated the mean difference in healthcare costs to be £5227 (95 % CrI £2019 to £8434). There was a substantial difference between the groups in the costs of informal care, with those in the LTPP + TAU group receiving less, but the difference was found to be quite uncertain (mean difference –£7613, 95 % CrI −£32,812 to £21,852). The cost of work days lost due to sickness was slightly lower for those in the TAU group, but differences between the groups were overall small and with considerable uncertainty (mean difference £330, 95 % CrI −£1154 to £1714).

3.4. Effectiveness

Fig. 4 shows the developments of CORE-6D utilities and HRSD-17 scores, respectively, over the study period. The two scores show a similar pattern, but the difference in treatment effectiveness between the therapy period and post-therapy appears less pronounced on the basis of the CORE-6D than the HRSD-17. After adjusting for baseline differences, the LTPP + TAU group had higher average QALYs (mean difference 0.16, 95 % CrI −0.01 to 0.33) and a lower – that is, less depressed – weighted average HRSD-17 score (mean difference −1.46, 95 % CrI −2.99 to 0.12). One patient in the TAU group died by suicide. On this basis, we estimated the average difference in suicide rates to be 0.5 % higher in the TAU group (95 % CrI −0.9 to 2.2).

3.5. Cost-effectiveness

Fig. 5 shows credible ellipses, which depict the uncertainty around the estimates when costs and effects are combined. They are pictorial versions of credible intervals. In this case, the credible ellipses are almost entirely contained in the north-east quadrant of the cost-effectiveness plane in the case of costs and QALYs, and almost entirely in the north-west quadrant with respect to costs and HRSD-17. This

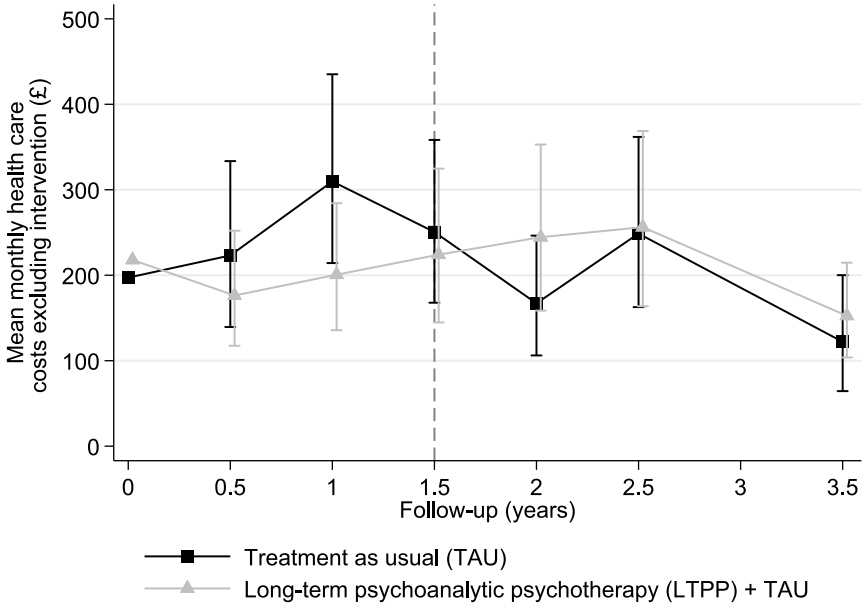

Fig. 3. Trajectory of non-intervention healthcare costs.

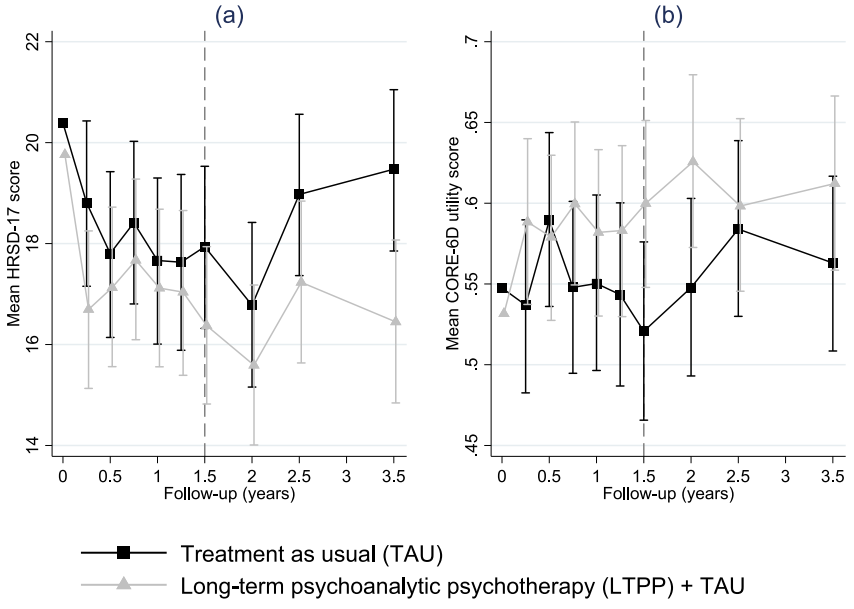

Fig. 4. Trajectory of HRSD-17 scores (a) and CORE-6D utility scores (b).

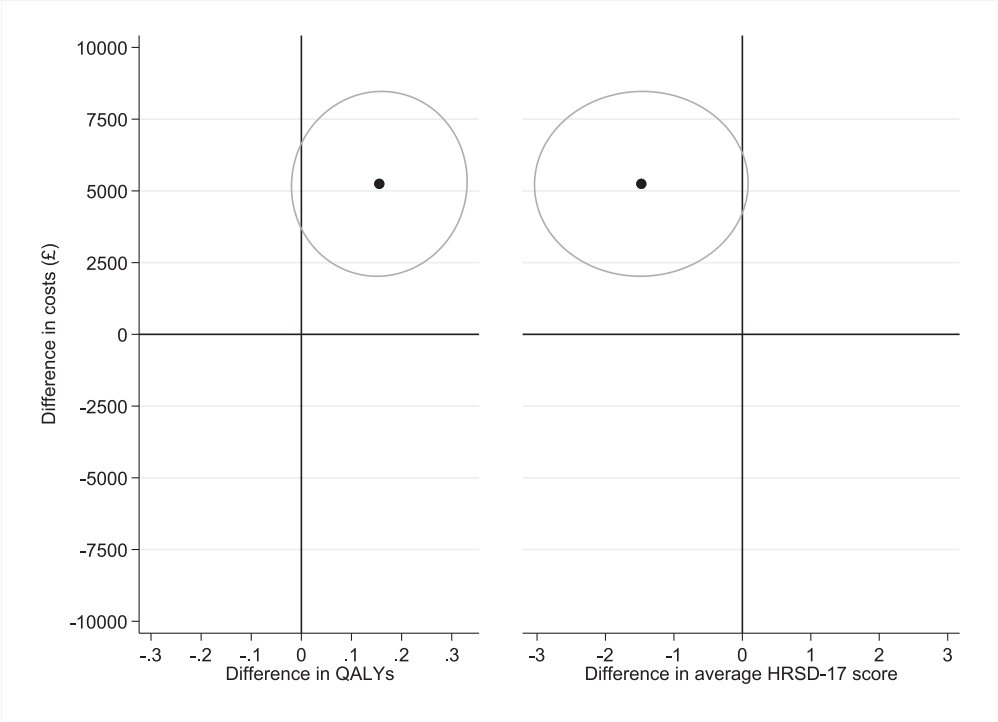

Fig. 5. Cost-effectiveness planes comparing outcomes in long-term psychodynamic psychotherapy augmentation of treatment as usual minus treatment as usual with 95 % credible ellipses (base case scenario).

indicates that LTPP + TAU had a probability of >90 % of being both more expensive and more effective than TAU (regardless of the outcome measure used). In the base case analysis, the ICER was approximately £33,000 per QALY. The probability that LTPP + TAU was cost-effective compared with TAU was 18 % at a value placed on a QALY gain of £20,000 (see Table 1 and Supplementary Fig. 9).

3.6. Sensitivity analyses

The results of the base case analysis were highly sensitive to variations in the cost of LTPP. Reducing its cost by 25 % decreased the ICER to ≈£24,500; raising it by 25 % increased the ICER to ≈£41,800. The alternative approach to costing service use had no noticeable effect on the estimated ICER, other than when only service use recorded in medical records was used; this reduced the ICER to ≈£27,700. However,

Table 1
Results of the primary analysis and sensitivity analyses.

Scenario	Parameter	Mean estimate			95 % credible interval of difference		Cost-effectiveness ratio (£/QALY)	Probability of being cost-effective at WTP of £20,000/QALY (%)
		TAU	LTPP + TAU	Difference	Lower bound	Upper bound		
I: Base case analysis	HC costs	8960	8724	5227	2019	8434	33,130	18
	QALYs	1.827	1.984	0.158	−0.013	0.33		
	HRSD	18.34	16.87	−1.46	−2.99	0.12		
IIa: Intervention cost decreased by 25 %	HC costs	8960	7358	3861	675	7021	24,472	38
IIb: Intervention cost increased by 25 %	HC costs	8960	10,089	6592	3344	9841	41,782	7
IIIa: Alternative combination of cost sources	HC costs	9453	9138	5152	1878	8419	32,754	20
IIIb: Service use based on CSRI only	HC costs	7290	7118	5295	3147	7640	33,378	16
IIIc: Service use based on medical records only	HC costs	5671	4548	4339	3113	5526	27,738	25
IV: Missing CORE-6D data lower than predicted	QALYs	1.713	1.872	0.159	−0.012	0.329	32,890	19
V: Log-normal cost distribution	HC costs	15,458	13,838	3846	−4860	11,928	24,583	42

HRSD: time-weighted average Hamilton Rating Scale for Depression; QALYs: quality-adjusted life years; HC costs: healthcare costs (£); WTP: willingness to pay; LTPP: long-term psychoanalytic psychotherapy; TAU: treatment as usual; CSRI: Client Services Receipt Inventory; CORE-6D: Clinical Outcomes Routine Evaluation – 6 Dimensions.

there was no reason to assume that the medical records were more accurate than self-report. The analysis in which CORE-6D was assumed to be MNAR had little impact on the estimated cost-effectiveness (ICER ≈£32,900). The log-normal model yielded a lower ICER of ≈£24,600/QALY than the gamma cost model in the base case analysis. The log-normal cost model had a slightly better statistical fit than the gamma cost-model according to the deviance information criterion (41 vs. 43) (Spiegelhalter et al., 2002). However, the posterior predictive checks for the log-normal distribution produced unrealistic replications of the observed data with a non-negligible proportion of patients predicted to have a cost of service use of more than £30,000 per observation period. Such values were not present in the observed data.

4. Discussion

The aim of this study was to assess the cost-effectiveness of LTPP + TAU compared with TAU for individuals with TRD. We found that LTPP + TAU had a probability of >90 % of being both more effective and more expensive than TAU, a difference that was sustained at the end of a 2-year follow-up after treatment and irrespective of the outcome measure used. More specifically, LTPP had a treatment cost of more than £5000 without substantial savings elsewhere compared with TAU provided in primary care settings in the UK. Its ICER was in the order of £33,000 per QALY gain. Of note, however, is that we found that this estimate was sensitive to changes in the cost of the intervention and to the statistical approach used to analyze the cost data.

Unfortunately, comparability to other health economic evaluations is limited not only in terms of their availability but also in terms of their heterogeneity (Shields and Elvidge, 2020). We are aware of four other RCTs of LTPP for depression: a Finnish study by Maljanen et al. (2016), two German studies, one by Huber et al. (2012) and the other by Leuzinger-Bohleber et al. (2018), and a Brazilian study conducted by Bastos et al. (2015). While patients in the TADS received on average 48 sessions of LTPP over a period of 18 months, treatment was substantially longer in all the other studies. For example, in Maljanen et al.'s (2016) study, participants received 232 sessions over a 5-year period and up to 2–3 sessions per week, and in Leuzinger-Bohleber et al.'s (2018) study participants received 234 sessions over 3 years. Another important difference was that the patients in the TADS were more severely and more protractedly depressed at baseline than those in the above RCTs: while the mean BDI score in the TADS was 36, it ranged between 18 and 32 in the other studies. Furthermore, 70 % of patients in Bastos et al.'s (2015) study were experiencing a first episode of depression rather than chronic depression or TRD. The range of failed previous treatment attempts in the TADS was between two and nine, with an average of 4.3. Again, this is much higher than those in the other studies, highlighting that the patients in the TADS had received available treatments recommended by NICE (2010) repeatedly without lasting effects.

To date, only Maljanen et al. (2016) report an economic evaluation alongside their RCT. As in our study, the differences they found in the costs of service use were due to the cost of the psychological therapy rather than to cost savings elsewhere. This contrasts with observational studies that have evaluated the economic impact of LTPP in non-TRD cohorts. Their results are generally more favorable with respect to LTPP (Berghout et al., 2010a,b; Beutel et al., 2004; de Maat et al., 2007).

4.1. Strengths and limitations

Although our study has a number of strengths, the results need to be interpreted in the context of several limitations as well. A particular strength is that the health economic data collection was integral to a pragmatic RCT, which had a well-defined sample of individuals with a diagnosis of TRD. A further strength is that the outcome ratings were not only reliable and rated by independent raters blind to participants' treatment allocation, but also included functioning and quality of life measures. This is important in particular as these outcomes have been reported by patients to be equally as important as remission of their depressive symptoms (Hummel et al., 2012). A further important strength is that the TAU control condition was a real-world comparison group. Therefore, the likelihood that the effects we estimated represent a causal relationship seems high. The same considerations suggest that our cost-effectiveness results might have high external validity. In terms of the study's limitations, we used a measure designed for depression research to calculate QALYs. Theoretically, this may increase the validity of the QALYs, but little is yet known about the properties of QALYs derived from the CORE-6D. Furthermore, although using two methods of collecting service use data and taking the higher value in the case of a discrepancy may have provided some remedy for the defects associated with using a solitary source, it was not possible to establish which of the two was the truer figure. Using the primary care reimbursement rates as the method to cost the LTPP provided face validity. However, the cost per session was at the upper end of the range often cited for cognitive-

behavioural therapy sessions, which might have implications for comparisons (Barrett and Petkova, 2013). However, the comparability, representativeness, and accuracy of such figures are notoriously difficult to establish among health economic studies. Although in the context of an RCT of the clinical effectiveness of long-term therapies the TADS sample size was adequate, typically, cost data are highly skewed (Mihaylova et al., 2011). Therefore, given the substantial amount of missing cost data, for an economic evaluation the same sample size was too small. This also had an impact in that we were unable to conduct credible subgroup analyses of responders/non-responders and alternative economic endpoints. A final limitation pertains to the generalizability of the results. In England and Wales, the provision of care for depression is of variable quality. The usual care for persisting forms of depression is often unacceptably suboptimal (Clark, 2011; Goyder et al., 2006; Wiles et al., 2018) Hence, the geographical generalizability of our results cannot be assumed.

5. Conclusion

Currently, NICE (2022) makes the following recommendations: (a) for a most plausible ICER below £20,000/QALY, health technologies will be assessed on cost-effectiveness and acceptability alone; (b) between £20,000 and £30,000/QALY, the ICER's uncertainty and aspects that relate to uncaptured benefits and non-health factors need to be considered; and (c) for ICERs of >£30,000/QALY, the case for factors other than cost-effectiveness needs to be increasingly strong. Our central estimate is that LTPP gives an ICER of £33,000/QALY gain, which falls above NICE's higher cost per QALY figure. This estimate is sensitive to the statistical approach used and the cost of the intervention, and the parameter uncertainty is large. However, given the trajectory of treatment benefits over the 2-year follow-up, there is reason to believe that the time horizon of this RCT was too short to capture all the benefits of LTPP. We did not capture the potential benefits of LTPP on patients' families and carers in this study, nor did the design allow the exploration of the impact of patients' treatment choice and treatment preference. Most importantly, there remains an unmet need for effective treatment options for TRD, and the provision of LTPP for TRD could aid in addressing health inequalities in this clinical area. Given the severity of TRD, it may be warranted to value each QALY gain in this population greater than in less severe conditions, but without undertaking calculations as described by NICE (2022) it is unclear whether TRD leads to a proportional or absolute QALY shortfall large enough to warrant a severity modifier. We believe that it is plausible that LTPP could affect mortality, for example, by influencing the risk of completed suicide. Obtaining reliable evidence on the magnitude and direction of this effect is challenging even in commonly used treatments for depression such as antidepressants (Nischal et al., 2012), and this study could not establish such an effect with a meaningful degree of certainty, but we believe that it is likely that the benefits of LTPP were somewhat underestimated because of this evidence gap. These considerations suggest areas for further research into LTPP and its possible cost-effectiveness. They include: (a) more detailed scrutiny of its direct costs; (b) further assessment of its longer-term effects; (c) the generalizability of an LTPP provision; and (d) identifying subgroups of patients for whom LTPP may be particularly helpful and cost-effective.

CRediT authorship contribution statement

LK, FR, & DT each contributed to writing this paper. FR and TB were responsible for data collection and preparation for analysis. LK carried out the statistical analyses. AG and PM provided advice on the analyses. PF, MK, PM & DG were responsible for the concept and design of the study. PF is the TADS Principal Investigator and has overall responsibility for its research management. DT was the TADS Clinical Director. All authors read and approved the final manuscript.

Declaration of competing interest

There are no financial conflicts of interest. In terms of non-financial interests, DT was responsible for writing and developing the TADS treatment manual describing the treatment method used.

Acknowledgements

The authors are indebted to the patients who participated in the TADS study. We would also like to thank the members of the Steering Committee and all previous researchers and trial managers, the research assistants, interns, students, administrators, and clinicians who worked on the trial. Finally, we thank all Primary Care centres for their help and support in getting access to the data.

Role of funding source

This research was funded by the Tavistock Clinic Charitable Foundation and the Tavistock & Portman NHS Foundation Trust. PF is in receipt of a National Institute for Health Research (NIHR) Senior Investigator Award (NF-SI-0514-10157). The views expressed are those of the authors and not necessarily those of the NHS, the NIHR or the Department of Health.

Appendix A. Supplementary material

Supplementary material to this article can be found online at https://doi.org/10.1016/j.jad.2023.04.109.

References

APA, 1994. Diagnostic and Statistical Manual of Mental Disorders: DSM-IV; Includes ICD-9-CM Codes Effective 1. Oct. 96. American Psychiatric Assoc.

Baio, G., 2012. Bayesian Methods in Health Economics. CRC Press.

Barrett, B., Petkova, H., 2013. Cognitive behaviour therapy: a comparison of costs. Unit Costs Health Soc. Care 16–20.

Bastos, A.G., Guimaraes, L.S.P., Trentini, C.M., 2015. The efficacy of long-term psychodynamic psychotherapy, fluoxetine and their combination in the outpatient treatment of depression. Psychother. Res. 25, 612–624. https://doi.org/10.1080/10503307.2014.935519.

Beck, A.T., Steer, R.A., Brown, G., 1996. Beck Depression Inventory–II. In: Psychol. Assess..

Beecham, J., Knapp, M., 2001. Costing psychiatric interventions. In: London, Thornicroft G. (Ed.), Gaskell In Measuring Mental Health Needs.

Berghout, C.C., Zevalkink, J., Hakkaart-van Roijen, L., 2010a. A cost-utility analysis of psychoanalysis versus psychoanalytic psychotherapy. Int. J. Technol. Assess. Health Care 26, 3–10. https://doi.org/10.1017/s0266462309990791.

Berghout, C.C., Zevalkink, J., Hakkaart-Van Roijen, L., 2010b. The effects of long-term psychoanalytic treatment on healthcare utilization and work impairment and their associated costs. J. Psychiatr Pr. 16, 209–216. https://doi.org/10.1097/01.pra.0000386907.99536.75.

Beutel, M.E., Rasting, M., Stuhr, U., Rüger, B., Leuzinger-Bohleber, M., 2004. Assessing the impact of psychoanalyses and long-term psychoanalytic therapies on health care utilization and costs. Psychother. Res. 14, 146–160. https://doi.org/10.1093/ptr/kph014.

Bovill, D., 2014. Patterns of Pay: Estimates From the Annual Survey of Hours and Earnings, UK, 1997 to 2013.

Clark, D.M., 2011. Implementing NICE guidelines for the psychological treatment of depression and anxiety disorders: the IAPT experience. Int. Rev. Psychiatry 23, 318–327. https://doi.org/10.3109/09540261.2011.606803.

Curtis, L., 2010. Unit Costs of Health And Social Care. University of Kent, Canterbury.

Curtis, L., Burns, A., 2018. Unit Costs of Health and Social Care 2018 | PSSRU. University of Kent.

DBIS, 2011. National Minimum Wage to rise from 1 October 2011 [WWW Document]. Natl. Minim. Wage Rise 1 Oct. 2011. https://www.gov.uk/government/news/national-minimum-wage-to-rise-from-1-october-2011 (accessed 9.18.22).

de Maat, S., Philipszoon, F., Schoevers, R., Dekker, J., De Jonghe, F., 2007. Costs and benefits of long-term psychoanalytic therapy: changes in health care use and work impairment. Harv. Rev. Psychiatry 15, 289–300. https://doi.org/10.1080/10673220701811654.

Edwards, S.J., Hamilton, V., Nherera, L., Trevor, N., 2013. Lithium or an atypical antipsychotic drug in the management of treatment-resistant depression: a systematic review and economic evaluation. Health Technol. Assess. Winch. Engl. 17, 1–190. https://doi.org/10.3310/hta17540.

Evans, J.M.-C., 2000. CORE: clinical outcomes in routine evaluation. J. Ment. Health 9, 247–255. https://doi.org/10.1080/jmh.9.3.247.255.

Fekadu, A., Donocik, J.G., Cleare, A.J., 2018. Standardisation framework for the Maudsley staging method for treatment resistance in depression. BMC Psychiatry 18, 100. https://doi.org/10.1186/s12888-018-1679-x.

Fenwick, E., O'Brien, B.J., Briggs, A., 2004. Cost-effectiveness acceptability curves – facts, fallacies and frequently asked questions. Health Econ. 13, 405–415. https://doi.org/10.1002/hec.903.

Fonagy, P., Rost, F., Carlyle, J., McPherson, S., Thomas, R., Pasco Fearon, R.M., Goldberg, D., Taylor, D., 2015. Pragmatic randomized controlled trial of long-term psychoanalytic psychotherapy for treatment-resistant depression: the Tavistock adult depression study (TADS). World Psychiatry 14, 312–321. https://doi.org/10.1002/wps.20267.

Goyder, E., Dibben, C., Grimsley, M., Peters, J., Blank, L., Ellis, E., 2006. Variation in prescribing for anxiety and depression: a reflection of health inequalities, cultural differences or variations in access to care? Int. J. Equity Health 5, 4. https://doi.org/10.1186/1475-9276-5-4.

Hamilton, M., 1960. A rating scale for depression. J. Neurol. Neurosurg. Psychiatry 23, 56.

Huber, D., Zimmermann, J., Henrich, G., Klug, G., 2012. Comparison of cognitive-behaviour therapy with psychoanalytic and psychodynamic therapy for depressed patients–a three-year follow-up study. Z. Für Psychosom. Med. Psychother. 58, 299–316.

Hummel, M.J.M., Volz, F., van Manen, J.G., Danner, M., Dintsios, C.-M., IJzerman, M.J., Gerber, A., 2012. Using the analytic hierarchy process to elicit patient preferences. Patient Patient-Centered Outcomes Res. 5, 225–237. https://doi.org/10.1007/BF03262495.

Jobst, A., Brakemeier, E.-L., Buchheim, A., Caspar, F., Cuijpers, P., Ebmeier, K.P., Falkai, P., Jan van der Gaag, R., Gaebel, W., Herpertz, S., Kurimay, T., Sabaß, L., Schnell, K., Schramm, E., Torrent, C., Wasserman, D., Wiersma, J., Padberg, F., 2016. European psychiatric association guidance on psychotherapy in chronic depression across Europe. Eur. Psychiatry 33, 18–36. https://doi.org/10.1016/j.eurpsy.2015.12.003.

Johnston, K.M., Powell, L.C., Anderson, I.M., Szabo, S., Cline, S., 2019. The burden of treatment-resistant depression: a systematic review of the economic and quality of life literature. J. Affect. Disord. 242, 195–210. https://doi.org/10.1016/j.jad.2018.06.045.

Lambert, P.C., Billingham, L.J., Cooper, N.J., Sutton, A.J., Abrams, K.R., 2008. Estimating the cost-effectiveness of an intervention in a clinical trial when partial cost information is available: a bayesian approach. Health Econ. 17, 67–81. https://doi.org/10.1002/hec.1243.

Leichsenring, F., Luyten, P., Hilsenroth, M.J., Abbass, A., Barber, J.P., Keefe, J.R., Leweke, F., Rabung, S., Steinert, C., 2015. Psychodynamic therapy meets evidence-based medicine: a systematic review using updated criteria. Lancet Psychiatry 2, 648–660. https://doi.org/10.1016/S2215-0366(15)00155-8.

Leuzinger-Bohleber, M., Hautzinger, M., Fiedler, G., Keller, W., Bahrke, U., Kallenbach, L., Kaufhold, J., Ernst, M., Negele, A., Schoett, M., Küchenhoff, H., Günther, F., Rüger, B., Beutel, M., 2018. Outcome of psychoanalytic and cognitive-behavioural long-term therapy with chronically depressed patients: a controlled trial with preferential and randomized allocation. Can. J. Psychiatr. https://doi.org/10.1177/0706743718780340, 0706743718780340.

Maljanen, T., Knekt, P., Lindfors, O., Virtala, E., Tillman, P., Härkänen, T., Helsinki Psychotherapy Study Group, 2016. The cost-effectiveness of short-term and long-term psychotherapy in the treatment of depressive and anxiety disorders during a 5-year follow-up. J. Affect. Disord. 190, 254–263. https://doi.org/10.1016/j.jad.2015.09.065.

Mavranezouli, I., Brazier, J.E., Rowen, D., Barkham, M., 2012. Estimating a preference-based index from the clinical outcomes in routine evaluation-outcome measure (CORE-OM) valuation of CORE-6D. Med. Decis. Mak. 33 (3), 381–395, 0272989X12464431.

McIntyre, R.S., Filteau, M.-J., Martin, L., Patry, S., Carvalho, A., Cha, D.S., Barakat, M., Miguelez, M., 2014. Treatment-resistant depression: definitions, review of the evidence, and algorithmic approach. J. Affect. Disord. 156, 1–7. https://doi.org/10.1016/j.jad.2013.10.043.

Mihaylova, B., Briggs, A., O'Hagan, A., Thompson, S.G., 2011. Review of statistical methods for analysing healthcare resources and costs. Health Econ. 20, 897–916. https://doi.org/10.1002/hec.1653.

Mrazek, D.A., Hornberger, J.C., Altar, C.A., Degtiar, I., 2014. A review of the clinical, economic, and societal burden of treatment-resistant depression: 1996–2013. Psychiatr. Serv. 65, 977–987. https://doi.org/10.1176/appi.ps.201300059.

Nemeroff, C.B., Heim, C.M., Thase, M.E., Klein, D.N., Rush, A.J., Schatzberg, A.F., Ninan, P.T., McCullough, J.P., Weiss, P.M., Dunner, D.L., Rothbaum, B.O., Kornstein, S., Keitner, G., Keller, M.B., 2003. Differential responses to psychotherapy versus pharmacotherapy in patients with chronic forms of major depression and childhood trauma. Proc. Natl. Acad. Sci. 100, 14293–14296. https://doi.org/10.1073/pnas.2336126100.

NICE, 2010. Depression - The NICE Guideline on the Treatment and Management of Depression in Adults: Updated Edition.

NICE, 2022. NICE Health Technology Evaluations: The Manual.

Nischal, Anil, Tripathi, A., Nischal, Anuradha, Trivedi, J.K., 2012. Suicide and antidepressants: what current evidence indicates. Mens Sana Monogr. 10, 33–44. https://doi.org/10.4103/0973-1229.87287.

Plummer, M., 2016. JAGS: Just Another Gibbs Sampler v. 4.0. 0.

Rost, F., Luyten, P., Fearon, P., Fonagy, P., 2019. Personality and outcome in individuals with treatment-resistant depression—Exploring differential treatment effects in the Tavistock Adult Depression Study (TADS). J. Consult. Clin. Psychol. 87, 433.

Ruhé, H.G., van Rooijen, G., Spijker, J., Peeters, F.P., Schene, A.H., 2012. Staging methods for treatment resistant depressionA systematic review. J. Affect. Disord. 137, 35–45.

Rush, A.J., Trivedi, M.H., Wisniewski, S.R., Nierenberg, A.A., Stewart, J.W., Warden, D., Niederehe, G., Thase, M.E., Lavori, P.W., Lebowitz, B.D., McGrath, P.J., Rosenbaum, J.F., Sackeim, H.A., Kupfer, D.J., Luther, J., Fava, M., 2006. Acute and longer-term outcomes in depressed outpatients requiring one or several treatment steps: a STAR*D report. Am. J. Psychiatry 163, 1905–1917. https://doi.org/10.1176/ajp.2006.163.11.1905.

Satyanarayana, S., Enns, M.W., Cox, B.J., Sareen, J., 2009. Prevalence and correlates of chronic depression in the Canadian community health survey: mental health and well-being. Can. J. Psychiatry Rev. Can. Psychiatr. 54, 389–398. https://doi.org/10.1177/070674370905400606.

Shields, G.E., Elvidge, J., 2020. Challenges in synthesising cost-effectiveness estimates. Syst. Rev. 9, 289. https://doi.org/10.1186/s13643-020-01536-x.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. J. RStat. Soc. Ser. B Stat. Methodol. 64, 583–639. https://doi.org/10.1111/1467-9868.00353.

Taylor, D., Carlyle, J., McPherson, S., Rost, F., Thomas, R., Fonagy, P., 2012. Tavistock adult depression study (TADS): a randomised controlled trial of psychoanalytic psychotherapy for treatment-resistant/treatment-refractory forms of depression. BMC Psychiatry 12, 60. https://doi.org/10.1186/1471-244X-12-60.

Thomas, L., Kessler, D., Campbell, J., Morrison, J., Peters, T.J., Williams, C., Lewis, G., Wiles, N., 2013. Prevalence of treatment-resistant depression in primary care: cross-sectional data. Br. J. Gen. Pract. 63, e852–e858. https://doi.org/10.3399/bjgp13X675430.

Whitehead, S.J., Ali, S., 2010. Health outcomes in economic evaluation: the QALY and utilities. Br. Med. Bull. 96, 5–21. https://doi.org/10.1093/bmb/ldq033.

Wiles, N., Taylor, A., Turner, N., Barnes, M., Campbell, J., Lewis, G., Morrison, J., Peters, T.J., Thomas, L., Turner, K., Kessler, D., 2018. Management of treatment-resistant depression in primary care: a mixed-methods study. Br. J. Gen. Pr. 68, e673–e681. https://doi.org/10.3399/bjgp18X699053.

Williams, J.W., 2001. Standardizing the Hamilton depression rating scale: past, present, and future. Eur. Arch. Psychiatry Clin. Neurosci. 251, 6–12. https://doi.org/10.1007/BF03035120.