European Psychiatry 57 (2019) 10–18

Contents lists available at ScienceDirect

European Psychiatry

journal homepage: http://www.europsy-journal.com

Original article

Cost-effectiveness analysis of collaborative treatment of late-life depression in primary care (GermanIMPACT)

Thomas Grochtdreis[a,*], Christian Brettschneider[a], Frederike Bjerregaard[b], Christiane Bleich[c], Sigrid Boczor[d], Martin Härter[c], Lars P. Hölzel[b,e], Michael Hüll[f,g], Thomas Kloppe[d], Wilhelm Niebling[h], Martin Scherer[d], Iris Tinsel[h], Hans-Helmut König[a]

[a] Department of Health Economics and Health Services Research, Hamburg Center for Health Economics, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[b] Department of Psychiatry and Psychotherapy, Medical Center – University of Freiburg, Faculty of Medicine, University of Freiburg, Freiburg, Germany
[c] Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[d] Department of General Practice / Primary Care, University Medical Center Hamburg-Eppendorf, Hamburg, Germany
[e] Parkklinik Wiesbaden Schlangenbad, Schlangenbad, Germany
[f] Center for Psychiatry Emmendingen, Emmendingen, Germany
[g] Center for Geriatric Medicine and Gerontology, Medical Center - University of Freiburg, Faculty of Medicine, Freiburg, Germany
[h] Division of General Practice, Medical Center - University of Freiburg, Faculty of Medicine, Freiburg, Germany

ARTICLE INFO

Article history:
Received 4 October 2018
Received in revised form 11 December 2018
Accepted 13 December 2018
Available online 15 January 2019

Keywords:
Depressive disorders
Late-life depression
Cost-effectiveness analysis
Depression-free days
Aged
Primary care
Collaborative care

ABSTRACT

Background: Late-life depression is a highly prevalent disorder that causes a large economic burden. A stepped collaborative care program was set up in order to improve care for patients with late-life depression in primary care in Germany: GermanIMPACT is the adaption of the Improving Mood-Promoting Access to Collaborative Treatment (IMPACT) program that has already been established in primary care in the USA. The aim of this study was to determine the cost-effectiveness of GermanIMPACT compared with treatment as usual from a societal perspective.
Methods: This study is part of a 12-month bi-centric cluster-randomized controlled trial aiming to assess the effectiveness of GermanIMPACT compared with treatment as usual among patients with late-life depression. A cost-effectiveness analysis using depression-free days (DFDs) was performed. Net-monetary benefit (NMB) regressions adjusted for baseline differences for different willingness-to-pay (WTP) thresholds were conducted and cost-effectiveness acceptability curves were constructed.
Results: In total, n = 246 patients (intervention group: n = 139; control group: n = 107) with a mean age of 71 from 71 primary care practices were included in the analysis. After 12 months, adjusted mean differences in costs and DFDs between intervention group and control group were +€354 and +21.4, respectively. Only the difference in DFDs was significant (p = 0.022). According to the unadjusted incremental cost-effectiveness ratio, GermanIMPACT was dominant compared with treatment as usual. The probability of GermanIMPACT being cost-effective was 80%, 90% or 95% if societal WTP per DFD was ≥€70, ≥€110 or ≥€180, respectively.
Conclusion: Evidence for cost-effectiveness of GermanIMPACT relative to treatment as usual is not clear. Only if societal WTP was ≥€180 for an additional DFD, GermanIMPACT could be considered cost-effective with certainty.

1. Introduction

Depressive disorders are highly prevalent in late-life populations and cause a large economic burden [1,2]. The point prevalence of major depression was estimated to be 7.2% (range 4.6%–9.3%) for individuals aged 75 years and older [1]. The total annual costs of major depression were €92 billion in Europe, of which €38 billion were associated with direct health care costs and €54 billion with indirect costs [3]. For Germany, health care costs of depressed individuals in late life were estimated to be approximately 30%–86% higher than health care costs of non-depressed individuals [2,4].

Depression in late life is associated with reduced quality of life, increased functional impairment as well as higher suicidality and

* Corresponding author at: Department of Health Economics and Health Services Research, Hamburg Center for Health Economics, University Medical Center Hamburg-Eppendorf, Martinistr. 52, 20246, Hamburg, Germany.
E-mail address: t.grochtdreis@uke.de (T. Grochtdreis).

http://dx.doi.org/10.1016/j.eurpsy.2018.12.007
0924-9338/

mortality [5–8]. Furthermore, late-life depression increases physical comorbidity and cognitive impairment [9]. This health burden influences health service use associated with late-life depression. According to a German study, individuals with late-life depression consulted their primary care physician (PCP) 8–10 times per year, and therefore more than twice as often as individuals without depression; yet, psychiatrists or psychotherapists were consulted only rarely [10].

In order to address the diversity and complexity of late-life depression in primary care, a collaborative care approach is recommended by national clinical guidelines in Europe [11,12]. Such a collaborative care approach for late-life depression is the Improving Mood-Promoting Access to Collaborative Treatment (IMPACT) program which has already been proven effective and cost-effective in the USA [13,14]. The IMPACT program has been recently adapted to the German primary care context and a compatible concept has been developed and proven effective [15,16].

As reported by a systematic review of the cost-effectiveness of collaborative care for the treatment of depressive disorders in primary care, three studies were conducted in late-life populations [17–20]. Two studies evaluated the cost-effectiveness of stepped collaborative care programs to prevent depression and anxiety in elderly home residents and primary care patients at risk for depression [17,19]. The third study evaluated the cost-effectiveness of a stepped collaborative care program for patients who were screened positive for untreated late-life depression in primary care [18]. Overall, the systematic review found ambiguous evidence on the cost-effectiveness of collaborative care depending on willingness-to-pay (WTP) [20]. To our knowledge, no study has evaluated the cost-effectiveness of health care networks for patients with late-life depression in Germany.

Therefore, the aim of this study was to determine the cost-effectiveness of a collaborative treatment for patients with late-life depression in German primary care (GermanIMPACT) compared with treatment as usual (TAU) from a societal perspective.

2. Material and methods

2.1. Sample

This study is part of a 12-month bi-centric cluster-randomized controlled trial to determine the effectiveness of a collaborative treatment for patients with late-life depression in German primary care (GermanIMPACT; German Clinical Trials Register identifier: DRKS00003589). The primary endpoint of this trial was the proportion of patients remitted of depressive symptoms. Study participants were recruited in randomized PCP practices around the centers of the German cities Hamburg, Freiburg and its region. Recruitment took place between July 2012 and September 2014. Inclusion criteria were an age of 60 years or older and showing moderate depressive symptoms in the scope of a depressive episode, recurrent depressive disorder, or dysthymia diagnosed by the PCP based on the International Classification of Diseases (ICD-10). Patients had to score between 10–14 points on the Patient Health Questionnaire-9 (PHQ-9) [21], indicating moderate symptoms of depressive disorder. Exclusion criteria were alcohol or drug abuse, severe cognitive impairment, severe psychological disorders, suicidal ideation as well as active depression treatment by a licensed psychotherapist.

Study participants were asked to complete pseudonymous questionnaires by mail. Data were assessed at begin of the intervention (baseline, T0), after 6 months (T1) and after 12 months (post-treatment, T2). All patients gave their written informed consent for study participation prior to the study. A detailed description of the study methods can be found elsewhere [15,16].

2.2. Intervention

The intervention group (IG) consisted of patients treated by PCP that were part of the stepped, collaborative, and coordinated health care network [16]. The health network was formed around the PCP together with a care manager and a consultant psychiatrist or licensed psychotherapist. The care manager was personally introduced to the patient by the PCP and on an eight-weekly basis, treatment evaluation sessions were held to facilitate treatment adaption by the PCP according to a stepped-care algorithm (e.g. medication dosage changes, adaption of care manager sessions) and in consultation with a psychiatrist, if necessary. The PCP was asked to treat depression in line with current clinical guidelines and to prescribe or change antidepressants, if necessary or if patients' depressive symptom severity did not (sufficiently) improve [12]. The evidence-based intervention consisted of a patient manual as well as an initial face-to-face session and ongoing telephone sessions between the care manager and the patient every other week. Care managers were trained nurses who received regular supervision by a consultant psychiatrist or licensed psychotherapist to discuss every patient's status. Every session consisted of monitoring of the intervention, medication and symptoms as well as planning and evaluation of behavioral activation tasks. Patients' depressive symptom severity was regularly assessed by the PHQ-9. Furthermore, sessions focusing on problem-solving techniques were optionally held to reduce depressive symptoms by teaching patients how to solve psychosocial problems systematically [22,23]. The intervention was based on a stepped care algorithm that allowed modifying the intervention every eight weeks by the care manager and the supervisor according to the patients' current mental health status.

The control group (CG) consisted of patients who attended regular PCP without involvement of a care manager and received TAU. Depressive symptom severity of patients in the CG was not routinely assessed by the PCP.

2.3. Measures

2.3.1. Patient Health Questionnaire-9

Depression severity was assessed retrospectively for a period of 2 weeks by the German version of the PHQ-9 [21,24–26]. It consists of nine items, each with four ordinal levels: 'not at all',' several days', 'more than half the days' and' nearly every day'. The PHQ-9 symptom severity score can range between 0 and 27, where a score between 5 and 9 defines mild depressive symptoms, a score between 10 and 14 defines moderate depressive symptoms and a score between 15 and above defines moderately severe to severe depressive symptoms [24].

2.3.2. EQ-5D-3L and EQ-5D-Index

Health-related quality of life was assessed by the German version of the EQ-5D-3L (hereafter EQ-5D) [27]. It consists of five dimensions: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. Each dimension is divided in three ordinal levels: 'no problems', 'moderate problems', and 'extreme problems' [27]. The EQ-5D has already been used successfully for late-life populations [28]. The construct validity and responsiveness was considered to be good [29]. For each of the 243 possible EQ-5D descriptive health states, an EQ-5D-Index value is available, with values ranging from 0 ("death") to 1 ("perfect health") [30]. To each patients' descriptive health state, an EQ-5D-Index value was assigned accordingly.

2.3.3. Service utilization and costs

Service utilization was assessed retrospectively over 6 months using the FIMA, a German questionnaire for health-related resource utilization in a late-life population [31]. Patients were asked about their utilization of outpatient physician (e.g. PCP,

https://doi.org/10.1016/j.eurpsy.2018.12.007 Published online by Cambridge University Press

specialist physician, psychotherapy) and non-physician services (e.g. physiotherapy, occupational therapy, massage), inpatient care, rehabilitation, as well as formal nursing care (provided by professional nurses or housekeepers) and informal nursing care (provided by family or friends). Furthermore, patients were asked about their use of medication and medical devices. Indirect costs due to absenteeism and reduced productivity at paid work were not assessed in the questionnaire because those costs are less relevant for late-life populations.

2.3.4. Other illness-related measures and sociodemographic variables

Symptom severity of comorbid generalized anxiety disorder was assessed by the Generalized Anxiety Disorder-7 (GAD-7) questionnaire [32]. Self-reported sociodemographic variables assessed at baseline were age, sex, statutory or private health insurance, marital status, living situation, level of education and employment status. Furthermore, co-morbid somatic diseases according to the Comorbidity Disease Index [33] were assessed.

2.4. Data analysis

2.4.1. Calculation of health effects

The primary measures of health effects for the cost-effectiveness analysis were depression-free days (DFDs). DFDs were estimated during the whole 12-month follow-up period and were based on the PHQ-9 symptom severity score at each measurement. A day in which the PHQ-9 symptom severity score was <5, was characterized as fully depression-free and a day, in which the PHQ-9 symptom severity score was ≥15, was characterized as a fully depressed day [24]. If a patient was considered depression-free (PHQ-9 < 5) at two consecutive measurements, the period between those measurements was considered also depression-free. If a patient had a PHQ-9 symptom severity score of ≥15 at two consecutive measurements, the period between those measurements was considered as fully depressed. To derive the number of DFDs between measurements in all other cases, linear interpolation of PHQ-9 symptom severity scores was used [34–36].

The secondary measure of health effects were quality-adjusted life years (QALYs). QALYs were calculated by weighting the duration of health states with the EQ-5D-Index values. Individual EQ-5D-Index values were linearly interpolated between measurements to derive QALYs over the 12-month follow-up period.

2.4.2. Calculation of costs

Costs were calculated in Euro (€) for the year 2013 from a societal perspective. Monetary valuation of each patients' direct physician and non-physician resource utilization was achieved using standardized unit costs [37]. Unit costs were inflated to 2013 price levels using the German gross domestic product inflation rate if no unit cost data for 2013 were available [38,39]. Monetary valuation of medication cost was achieved using the pharmacy retail price of the German official pharmaceutical index [40].

Intervention costs consisted of the labor costs for the care manager who performed the face-to-face and telephone sessions. Furthermore, intervention costs consisted of labor costs for the psychiatrist or licensed psychotherapist who supervised care managers every one to two weeks. Labor costs of nurses, psychiatrists and licensed psychotherapists were calculated based on data from the German official earnings and labor cost index [41,42].

2.4.3. Statistical analysis

The percentage of missing information across the 129 variables varied between 0.40% and 18.55%. In total, 41.94% of all patients had missing information. Incomplete variables were imputed using multiple imputation by chained equations (MICE) with predictive mean matching as imputation method [43]. Differences in baseline characteristics between the IG and CG were assessed using simple linear regression (F-tests). Differences in costs and effects between the IG and CG were calculated using multilevel mixed-effects linear regression with PCP practices as random effect. To analyze the cost-effectiveness of GermanIMPACT compared with TAU, the incremental cost-effectiveness ratio (ICER) was calculated as the ratio of the difference in mean cost and the difference in mean health effects (DFD or QALY) between IG and CG during the 12-month follow-up period. The ICER is a point estimate that compares the additional costs with the additional effects of an intervention over control [44]. Bootstrapped confidence intervals around the difference in mean cost and the difference in mean health effects between IG and CG have been calculated with seemingly unrelated regression [45].

Furthermore, in order to consider uncertainty in the data, cost-effectiveness acceptability curves (CEACs), based on the incremental net-monetary benefit (NMB) were constructed. The incremental NMB is a combination of costs and health effects, where health effects are transformed into monetary units. Thereby, a hypothesized maximum WTP value of a decision maker for an additional health effect is used as exchange rate [46]:

$$NMB_i = WTP * Effects_i - Costs_i$$

A CEAC plots the proportion of the sampling distribution of incremental costs and health effects that lie below the maximum WTP for a health effect gained [47].

The incremental NMB [44] of GermanIMPACT was calculated by multilevel mixed-effects linear regression adjusted for age, sex, total costs, EQ-5D-Index value, PHQ-9-Index value, comorbid neurologic diseases and cancer at baseline and with PCP practices as random effect. Based on the incremental NMB for different WTP thresholds, CEACs were constructed to plot the probability of cost-effectiveness of GermanIMPACT [44]. Multiple imputation and data analysis were performed using Stata/SE 14.1 (StataCorp, TX, USA). All the applied statistics were two-sided with a significance level of $p < 0.05$.

2.4.4. Additional analyses

First, analyses were repeated from a health care perspective by excluding informal nursing care costs. Second, analyses were repeated considering only mental health-related costs (psychiatric inpatient care, outpatient services by psychiatrists and psychotherapist, psychiatric medication). Last, analyses were repeated only with patients who completed questionnaires at all time-points of assessment (completers) as well as with cost-outliers winsorized at the 99th percentile.

3. Results

In total, n = 1963 PCP practices (n = 1116 in Hamburg and n = 847 in Freiburg) were assessed for eligibility to include patients to the study. Of those, 99 PCP practices were randomized to either IG or CG. PCP practices randomized to the IG identified n = 788 patients and PCP practices randomized to the CG identified n = 717 patients. Finally, n = 248 patients (IG: n = 139; CG: n = 109) from n = 71 PCP practices (IG: n = 39; CG: n = 32) were analyzed (Fig. 1). Two patients with total costs 11 and 16 times higher than the mean total costs of the remaining sample were excluded ex post from the analyses.

3.1. Baseline characteristics

Characteristics of the sample at baseline (IG: n = 139; CG: n = 107) are presented in Table 1. On average, patients were 71 years old and 78% were female, only 11% were still working and 62% were living with a partner/family. The mean PHQ-9-Index value was 10.2, the

https://doi.org/10.1016/j.eurpsy.2018.12.007 Published online by Cambridge University Press

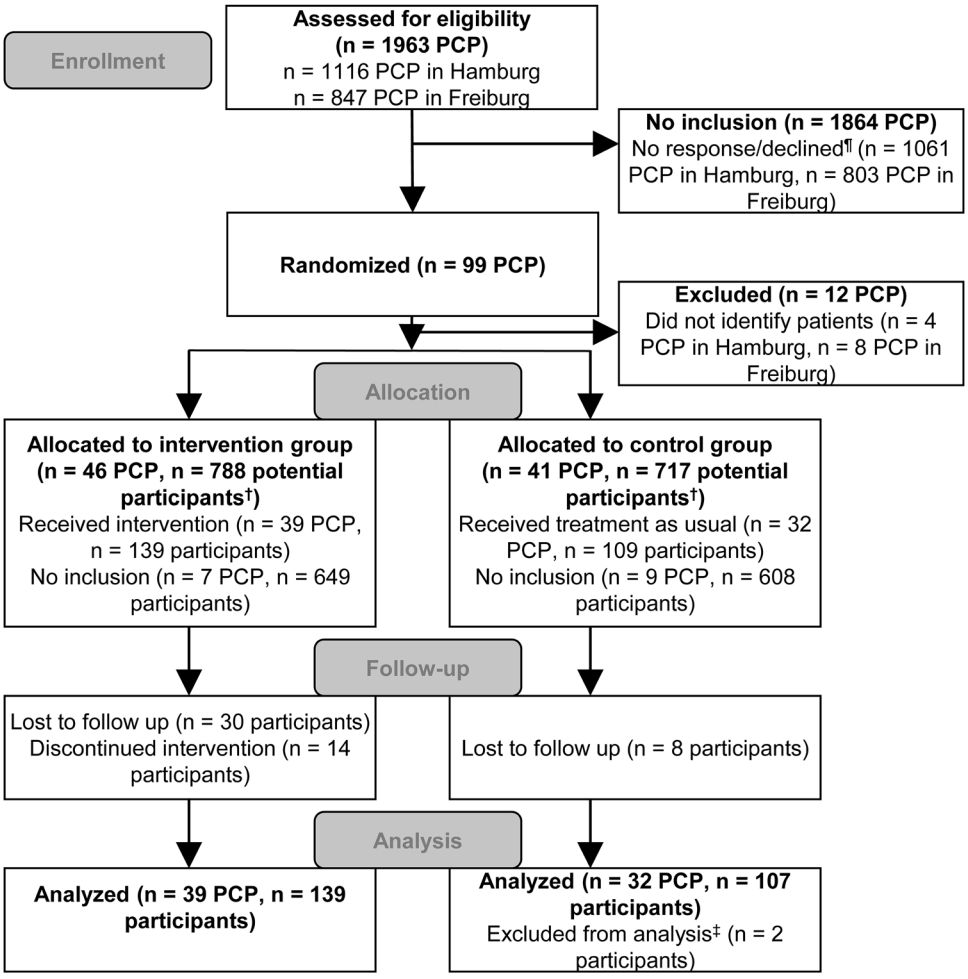

Fig. 1. CONSORT diagram showing the flow of clusters and participants.
PCP: Primary care physician
¶Mainly PCP did not respond to study invitation; PCP declined due to lack of interest or lack of time
†Potential participants with a diagnosed depressive episode, recurrent depressive disorder or dysthymia
‡Excluded patients had total costs 11 and 16 times higher than the mean total costs of the remaining sample.

mean GAD-7-Index value was 8.2 and the mean EQ-5D-Index value was 0.55. Only the mean PHQ-9-Index values were significantly different between IG (10.7) and CG (9.6; $p = 0.040$).

3.2. Effects and costs

During the 12-month follow-up period, DFDs were 207.1 in the IG and 185.8 in the CG. The adjusted difference in mean DFD was statistically significant (+21.4; $p = 0.022$). QALYs were 0.57 in the IG and 0.56 in the CG. The adjusted difference in QALYs was not statistically significant (+0.01; $p = 0.701$; Table 2).

Mean total costs during the 12-month follow-up period were €6105 in the IG and €6415 in the CG. Mean total intervention costs that were part of the mean total costs in the IG accrued to €164. Of those, the mean costs for the care manager were €123 and the mean costs for the supervision were €41, respectively. The adjusted differences in the respective cost categories and the adjusted difference in mean total costs (+€354; $p = 0.705$) were not significant (Table 2).

3.3. Cost-effectiveness

Both, the unadjusted ICER for an additional DFD and for an additional QALY showed dominance (less costs, more health effects) of the IG compared with the CG. The unadjusted mean total costs were significantly lower in the IG (−€314; $p = 0.049$, 95% CI −€627 to −€2) and the unadjusted mean DFD (+6.45; $p = 0.588$; 95% CI −16.92 to 29.83) and QALY (+0.00; $p = 0.999$; 95% CI −0.07 to 0.07) were both not significantly different between IG and CG.

The adjusted probability for cost-effectiveness of GermanIMPACT at a WTP of €0 per additional DFD/QALY was 35%. The adjusted probability for cost-effectiveness was 80%, 90% or 95% if WTP per additional DFD was ≥€70, ≥€110 or ≥€180, respectively (Fig. 2). For a WTP of €50,000 per additional QALY, the adjusted probability for cost-effectiveness was 50% (Fig. 3).

3.4. Additional analysis

From a health care perspective, the adjusted probability for cost-effectiveness was 95% if WTP per additional DFD was ≥€200. For a WTP of €50,000 per additional QALY, the adjusted probability for cost-effectiveness was 45%. When considering only mental health care costs, the adjusted probability for cost-effectiveness was 95% if WTP per additional DFD was ≥€10. For a WTP of €50,000 per additional QALY, the adjusted probability for cost-effectiveness was 80% (Table 3).

In the analysis considering completers (n = 202), the adjusted probability for cost-effectiveness was 95% if WTP per additional

https://doi.org/10.1016/j.eurpsy.2018.12.007 Published online by Cambridge University Press

Table 1
Comparison of sample characteristics at baseline (n = 246).

Characteristics	Intervention group (n = 139)	Control group (n = 107)	P-value‡
Age: mean years (SD)	71.13 (7.17)	71.60 (8.13)	0.632
Female sex: n (%)	107 (77.0)	85 (79.4)	0.644
Statutory health insurance: n (%)	132 (95.0)	103 (96.3)	0.627
Marital status: n (%)			0.168
Single/divorced/widowed	76 (54.7)	48 (45.3)	
Married/in partnership	64 (45.8)	58 (54.2)	
Living situation: n (%)			0.095
With partner/family	65 (46.8)	64 (59.8)	
Institutionalized	4 (2.9)	4 (3.7)	
Alone	70 (50.4)	39 (36.4)	
Education: n (%)			0.318
No graduation	4 (2.9)	2 (1.9)	
Low qualification	72 (51.8)	65 (60.7)	
Middle qualification	33 (23.7)	21 (19.6)	
High qualification	13 (9.4)	13 (12.1)	
University degree	17 (12.2)	6 (5.6)	
Employment: n (%)			0.878
Working	15 (10.8)	11 (10.3)	
Not employed	25 (18.0)	22 (20.6)	
Retired	99 (71.2)	74 (69.2)	
Co-morbidity domains: n (%)			
Neurologic	15 (10.8)	9 (8.4)	0.534
Cancer	15 (10.8)	21 (19.6)	0.055
EQ-5D-Index: mean (SD)	0.55 (0.31)	0.55 (0.31)	0.864
EQ-VAS: mean (SD)	57.75 (18.76)	59.69 (18.42)	0.421
PHQ-9-Index: mean (SD)¶	10.67 (4.02)	9.64 (3.62)	0.040*
GAD-7-Index: mean (SD)	8.53 (4.39)	7.71 (4.70)	0.153
Total costs: Mean (SD)	€2920 (€4425)	€4222 (€7729)	0.084

SD: Standard deviation, PHQ: Patient Health Questionnaire, GAD: Generalized Anxiety Disorder.
¶ Differences to Hölzel et al. [15] are due to different imputation strategies and baseline samples.
‡ Based on F-test.
* $p \leq 0.05$.

DFD was ≥€120. For a WTP of €50,000 per additional QALY, the adjusted probability for cost-effectiveness was 56%. In the analysis with winsorized cost-outliers (n = 248), the adjusted probability for cost-effectiveness was 95% if WTP per additional DFD was ≥€80. For a WTP of €50,000 per additional QALY, the adjusted probability for cost-effectiveness was 63% (Table 3).

4. Discussion

The aim of this study was to determine the cost-effectiveness of a collaborative treatment for patients with late-life depression in German primary care (GermanIMPACT). According to the results of this study, the collaborative treatment was effective relative to TAU

Table 2
Adjusted¶ differences between intervention group and control group in mean costs (by cost category) and health effects during 12-month follow up (n = 246).

Cost category / Measure of health effect	IG (n = 139)	CG (n = 107)	Diff. IG-CG	P-value‡
Intervention costs[1]: Mean (SE)	€164 (€8)	€0 (€0)		<0.001***
Inpatient care: Mean (SE)	€1659 (€359)	€1646 (€413)	€13 (€546)	0.982
Somatic inpatient care: Mean (SE)	€1654 (€353)	€1351 (€314)	€303 (€463)	0.512
Psychiatric inpatient care: Mean (SE)	€3 (€37)	€294 (€149)	−€291 (€160)	0.070
Rehabilitation: Mean (SE)	€398 (€1265)	€406 (€157)	−€7 (€202)	0.971
Outpatient physician services: Mean (SE)	€935 (€82)	€1069 (€81)	−€133 (€117)	0.253
Outpatient non-physician services: Mean (SE)	€1470 (€166)	€1125 (€153)	€343 (€229)	0.134
Medication: Mean (SE)	€1250 (€239)	€1154 (€214)	€98 (€316)	0.756
Somatic medication: Mean (SE)	€1096 (€237)	€1036 (214)	€60 (€313)	0.847
Psychiatric medication: Mean (SE)	€155 (€18)	€117 (€12)	€38 (€22)	0.082
Medical devices: Mean (SE)	€128 (€33)	€100 (€29)	€29 (€45)	0.524
Nursing care: Mean (SE)	€352 (€92)	€630 (€137)	−€278 (€154)	0.072
Formal nursing care: Mean (SE)	€111 (€26)	€174 (€36)	−€63 (€44)	0.154
Informal nursing care: Mean (SE)	€241 (€74)	€456 (€127)	−€215 (€141)	0.128
Total costs: Mean (SE)	€6395 (€583)	€6041 (€714)	€354 (€934)	0.705
Total costs from HCP[2]: Mean (SE)	€6155 (€555)	€5 (€686)	€558 (€893)	0.532
Total mental health care costs[3]: Mean (SE)	€321 (€80)	€856 (€326)	−€535 (€350)	0.126
QALY: Mean (SE)	0.57 (0.02)	0.56 (0.02)	0.01 (0.02)	0.701
DFD: Mean (SE)	207.12 (6.73)	185.84 (6.58)	21.38 (9.04)	0.022*

SE: Standard error, HCP: Health care perspective, QALY: Quality-adjusted life year, DFD: Depression-free day, IG: Intervention group, CG: Control group.
¶ Adjusted for costs, age, sex, PHQ-9-Index, EQ-5D-Index, comorbid domains neurologic and cancer at baseline by mixed-effects linear regression with robust standard errors.
[1] Unadjusted.
[2] Excluded cost categories: Informal nursing care.
[3] Included cost categories: Inpatient-care psychiatry, outpatient psychiatrist, outpatient psychotherapist, psychiatric medication.
‡ Based on F-test.
* $p \leq 0.05$.
*** $p \leq 0.001$.

https://doi.org/10.1016/j.eurpsy.2018.12.007 Published online by Cambridge University Press

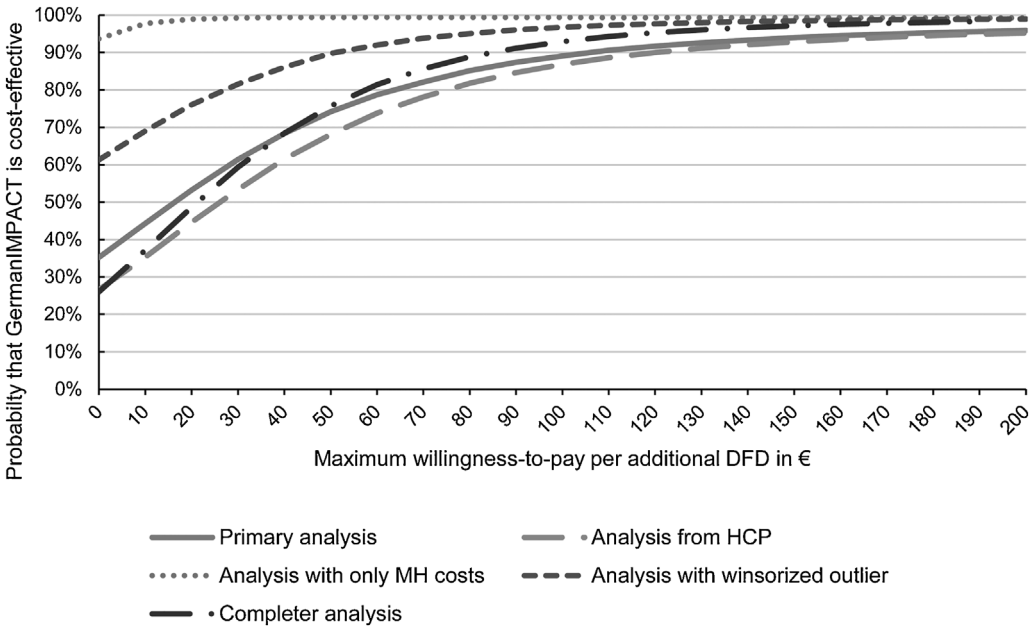

Fig. 2. Adjusted¶ cost-effectiveness acceptability curves for an additional DFD: Primary analysis and additional analyses for different costs and different cases.
DFD: Depression-free day, HCP: Health care perspective, MH: Mental health
¶Adjusted for costs, age, sex, EQ-5D-Index, PHQ-9-Index, comorbid domains neurologic and cancer at baseline by mixed-effects linear regression with robust standard errors.

by means of a significant increase in DFD, whereas the difference in societal costs was not significant. However, GermanIMPACT did not prove to be cost-effective with certainty (in terms of a 95% probability), unless societal WTP for an additional DFD was ≥€180. This was mainly due to the large statistical variance of incremental costs.

Even though the study was able to show a significant increase in DFD attributable to the collaborative treatment, this was not reflected by a simultaneous improvement in QALYs. This missing relationship is counterintuitive to studies that were able to show that EQ-5D-Index values were associated with clinical and quality of life improvement or deterioration in patients with depression [48]. However, studies that aimed to predict generic preference-based measures from psychometric measures for depression severity (e.g. PHQ-9) were able to explain only 31% and 48% of variance with their models [48,49]. This relatively low correlation might be interpreted as indicator for a limited relationship between depressiveness and health-related quality of life in the sample of this study. However, the current study was not powered to detect such a relationship.

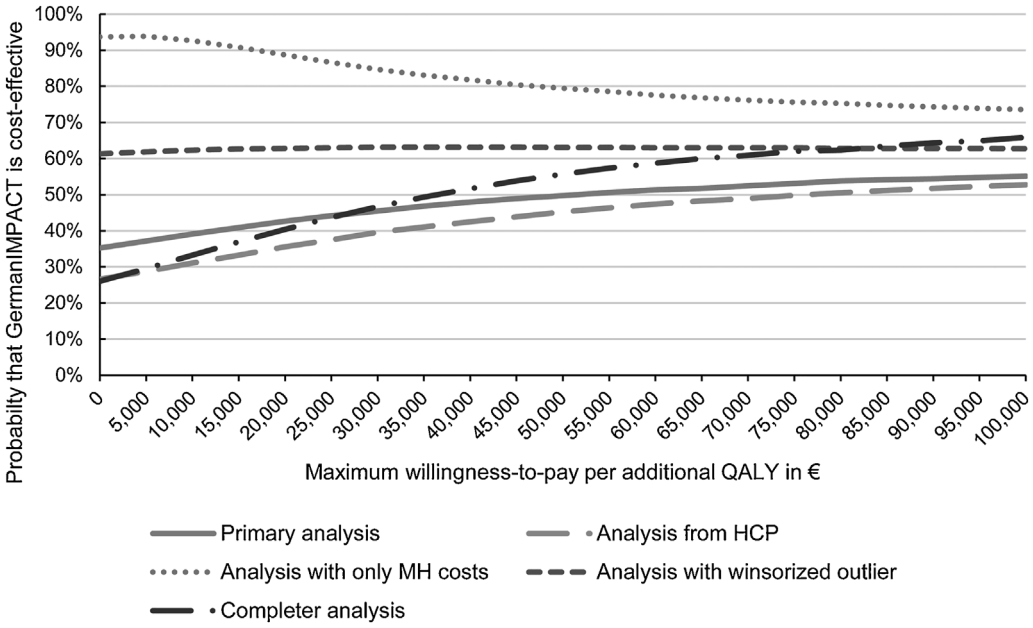

Fig. 3. Adjusted¶ cost-effectiveness acceptability curves for an additional QALY: Primary analysis and additional analyses for different costs and different cases.
QALY: Quality-adjusted life year, HCP: Health care perspective, MH: Mental health
¶Adjusted for costs, age, sex, EQ-5D-Index, PHQ-9-Index, comorbid domains neurologic and cancer at baseline by mixed-effects linear regression with robust standard errors.

https://doi.org/10.1016/j.eurpsy.2018.12.007 Published online by Cambridge University Press

Table 3
Additional analyses: Adjusted[¶] differences in costs and health effects between intervention group and control group during 12-month follow-up.

Analysis (Difference CG-IG)	N	Adj. costs (SE)	Adj. ΔQALY (SE)	Adj. ΔDFD (SE)
Primary analysis[1]	246	€354 (€933)	0.01 (0.02)	21.38 (9.35)*
Analysis from HCP[2]	246	€441 (€875)	0.01 (0.02)	21.38 (9.35)*
Analysis with only mental health costs[3]	246	−€553 (€354)	0.01 (0.02)	21.38 (9.35)*
Analysis with winsorized cost-outliers[4]	248	−€305 (€1058)	0.01 (0.02)	22.11 (8.89)*
Completer analysis	202	€615 (€957)	0.02 (0.02)	30.22 (9.85)**

SE: Standard error, CG: Control group, IG: Intervention group, QALY: Quality-adjusted life year, HCP: Health care perspective, DFD: Depression-free day.
¶ Adjusted for costs, age, gender, EQ-5D-Index, PHQ-9-Index, comorbid domains neurologic and cancer at baseline by mixed-effects linear regression with robust standard errors.
[1] Included cost categories: Intervention, inpatient care, rehabilitation, outpatient physician services, outpatient non-physician services, somatic and psychiatric medication, medical devices, and formal/informal nursing care.
[2] Excluded cost categories: Informal nursing care.
[3] Included cost categories: Inpatient-care psychiatry, outpatient psychiatrist, outpatient psychotherapist, psychiatric medication.
[4] Patients with total costs greater than the 95th percentile of the total costs of the particular group.
* $p \leq 0.05$.
** $p \leq 0.01$.

Based on the unadjusted ICER, GermanIMPACT was dominant compared with TAU. However, as this analysis was based on a bi-centric cluster-randomized controlled trial and the sample size was relatively small, results adjusted for imbalances in baseline characteristics are to be considered less prone to biased cost-effectiveness estimates [44]. Therefore, to handle uncertainty in the cost-effectiveness analysis, the net-benefit regression approach has been used for estimation enabling to control for imbalances [50]. In the study evaluating the cost-effectiveness of the IMPACT program for late-life primary care patients with depression in the USA, an ICER of US$60 per additional DFD was calculated. As GermanIMPACT was an adaption of the collaborative care program of that study, comparability between the IMPACT program and the program evaluated in the current study is given [13,15,16]. Furthermore, participants of that study were comparable by means of gender distribution and age. However, uncertainty of the ICER has not been quantified for the IMPACT program and none of the studies found clear evidence of cost-effectiveness for the collaborative treatments.

Moreover, whether the implementation of GermanIMPACT would represent an efficient use of resources remains unclear because there is no generally accepted WTP threshold value for an additional DFD. Unützer et al. [51] estimated that depressed primary care patients were willing to pay US$18 for an additional DFD. Our study revealed that the probability for cost-effectiveness at a WTP of US$18 would be only 49%.

Another study, evaluating the cost-effectiveness of a collaborative care program to prevent depression and anxiety in residents in homes for elderly in the Netherlands, calculated an ICER of US $34,755 per QALY [17]. The probability of cost-effectiveness at a WTP of >€50,000 for an additional QALY was only 46%. Because participants of that study were somewhat older and the aim of that study was prevention of depression and anxiety, comparability with the current study intervention is limited. However, in both studies, the collaborative care intervention was unlikely to be cost-effective, even for high WTP for an additional QALY.

4.1. Strengths and limitations

The first strength of this study is that the evaluated collaborative care program was adapted from a program that has already been proven effective and cost-effective in the USA and that it has been tailored to German primary care context. Second, for the assessment of service utilization, a questionnaire has been used that has been explicitly developed for late-life populations [31]. Third, multiple imputation has been used to handle missing information. Multiple imputation is superior to simple imputation methods because uncertainty of the predicted missing values is taken into account and distributions and relationships in the data are preserved [52]. Furthermore, compared with complete case analysis, the analysis with multiply imputed data is not considered to be biased under the missing at random assumption [53]. Last, the NMB for different WTP thresholds were estimated using covariates at baseline for adjustment. Yet, by using covariate-adjustment, possible bias from those covariates was removed.

However, the study has some limitations. First, in order to detect a rate of remission of depressive symptoms with adequate power, the required sample size of the GermanIMPACT trial was 250 patients. However, this sample size might have been too small to measure significant differences in costs, which showed, moreover, a large statistical variance. Second, patients in the IG had significantly more severe depressive symptoms at baseline compared with patients in the CG. PCP in the IG might have had increased interest in including patients with poorer health into the study than PCP in the CG, as they might expected relief in workload through the care manager. Furthermore, PCP in the CG might not have included patients with more severe depressive symptoms, as no additional intervention beyond TAU was provided. Third, generalizability of the results of this study to the population with late-life depression in Germany might be limited because study participants were recruited in randomized PCP practices around two study centers only. Last, out of 1505 patients identified potentially eligible, merely 18% were included in this study. This low participation rate might be explained by heterogeneous recruiting success by PCPs, non-existent acute depressive symptoms or a limited awareness of the disorder by late-life primary care patients, their motivation and believe in change and concerns about time-consuming recruitment processes. Furthermore, even though the PCP and the study coordinators decided about eligibility of study participation based on diagnosis of a depressive disorder and the patients' depressive symptom severity, patients decided themselves to participate or to refuse participation in the study. Thus, participation might have been influenced by non-acceptance stigma of depressive disorders by patients. Health services research is regularly challenged by such low participation rates and recruitment rates per PCP of below five patients as in the current study are common, nevertheless a selection bias cannot be ruled out [14,54,55].

5. Conclusion

There is no clear evidence for cost-effectiveness of the analyzed collaborative treatment for patients with late-life depression in German primary care relative to TAU. Only if societal WTP was

https://doi.org/10.1016/j.eurpsy.2018.12.007 Published online by Cambridge University Press

≥€180 for an additional DFD, GermanIMPACT could be considered cost-effective with certainty. The probability of cost-effectiveness of GermanIMPACT for an additional QALY was low, even for high WTP ≥€50,000. Further research is needed in order to analyze the cost-effectiveness of collaborative treatment for patients below 60 years of age with depression in German primary care. In order to reduce uncertainty in cost-effectiveness analyses of collaborative treatment, future research needs to be conducted with large sample sizes powered to detect differences in costs and health-related quality of life. Furthermore, future research needs to focus on the relationship between depressive symptom severity and cost-effectiveness of collaborative treatment in order to determine most efficient use of resources.

Financial or other support

This work was supported by the German Federal Ministry of Education and Research (grant number: 01GY1142).

Conflict of interest

The authors report no conflict of interest.

Acknowledgements

We thank Moritz Hadwiger, Angela Kotterer, Annemarie Wegener and Katharina Hussain for their support in dealing with data on medications, Judith Dams for her helpful comments on an earlier version of the manuscript and all involved patients and therapists for their participation.

References

[1] Luppa M, Sikorski C, Luck T, Ehreke L, Konnopka A, Wiese B, et al. Age- and gender-specific prevalence of depression in latest-life - systematic review and meta-analysis. J Affect Disord 2012;136(3):212–21.
[2] Bock JO, Brettschneider C, Weyerer S, Werle J, Wagner M, Maier W, et al. Excess health care costs of late-life depression - results of the AgeMooDe study. J Affect Disord 2016;199:139–47.
[3] 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.
[4] Luppa M, Heinrich S, Angermeyer MC, König H-H, Riedel-Heller SG. Healthcare costs associated with recognized and unrecognized depression in old age. Int Psychogeriatr 2008;20(6):1219–29.
[5] Beekman AT, Deeg DJ, Braam AW, Smit JH, Van Tilburg W. Consequences of major and minor depression in later life: a study of disability, well-being and service utilization. Psychol Med 1997;27(6):1397–409.
[6] Schulz R, Drayer RA, Rollman BL. Depression as a risk factor for non-suicide mortality in the elderly. Biol Psychiatry 2002;52(3):205–25.
[7] Unützer J, Patrick DL, Diehr P, Simon G, Grembowski D, Katon W. Quality adjusted life years in older adults with depressive symptoms and chronic medical disorders. Int Psychogeriatr 2000;12(1):15–33.
[8] Waern M, Runeson BS, Allebeck P, Beskow J, Rubenowitz E, Skoog I, et al. Mental disorder in elderly suicides: a case-control study. Am J Psychiatry 2002;159(3):450–5.
[9] Blazer DG. Depression in late life: review and commentary. J Gerontol A Biol Sci Med Sci 2003;58(3):249–65.
[10] Glaesmer H, Gunzelmann T, Martin A, Brahler E, Rief W. Die Bedeutung psychischer Beschwerden für die medizinische Inanspruchnahme und das Krankheitsverhalten Älterer. Psychiat Prax 2008;35(4):187–93.
[11] NICE. Depression in adults. The treatment and management of depression in adults. London (UK): National Institute for Health and Clinical Excellence; 2009.
[12] DGPPN, BÄK, KBV, AWMF, AkdÄ, BPtK, et al. S3-Leitlinie/Nationale VersorgungsLeitlinie Unipolareunipolare Depression - langfassung. Berlin, Düsseldorf: Leitliniengruppe Unipolare Depression; 2015.
[13] Katon WJ, Schoenbaum M, Fan M-Y, Callahan CM, Williams Jr. J, Hunkeler E, et al. The IMPACT investigators. Cost-effectiveness of improving primary care treatment of late-life depression. Arch Gen Psychiatry 2005;62(12):1313–20.
[14] Unützer J, Katon W, Callahan CM, Williams Jr. J, Hunkeler E, Harpole L, et al. The IMPACT investigators. Collaborative care management of late-life depression in the primary care setting: a randomized controlled trial. JAMA 2002;288 (22):2836–45.
[15] Hölzel LP, Bjerregaard F, Tinsel I, Kloppe T, Boczor S, Bleich C, et al. Coordinated treatment of depression in elderly people in primary care. A cluster-randomized, controlled study (GermanIMPACT). Dtsch Arztebl Int 2018;115 (44):741–7.
[16] Wernher I, Bjerregaard F, Tinsel I, Bleich C, Boczor S, Kloppe T, et al. Collaborative treatment of late-life depression in primary care (GermanIMPACT): study protocol of a cluster-randomized controlled trial. Trials 2014;15:351.
[17] Bosmans JE, Dozeman E, van Marwijk HWJ, van Schaik DJF, Stek ML, Beekman ATF, et al. Cost-effectiveness of a stepped care programme to prevent depression and anxiety in residents in homes for the older people: a randomised controlled trial. Int J Geriatr Psychiatry 2014;29(2):182–90.
[18] van der Weele GM, de Waal MWM, van den Hout WB, de Craen AJM, Spinhoven P, Stijnen T, et al. Effects of a stepped-care intervention programme among older subjects who screened positive for depressive symptoms in general practice: the PROMODE randomised controlled trial. Age Ageing 2012;41 (4):482–8.
[19] van't Veer-Tazelaar P, Smit F, van Hout H, van Oppen P, van der Horst H, Beekman A, et al. Cost-effectiveness of a stepped care intervention to prevent depression and anxiety in late life: randomised trial. BJP 2010;196(4):319–25.
[20] Grochtdreis T, Brettschneider C, Wegener A, Watzke B, Riedel-Heller S, Härter M, et al. Cost-effectiveness of collaborative care for the treatment of depressive disorders in primary care: a systematic review. PLoS One 2015;10(5) e0123078.
[21] Löwe B, Spitzer RL, Zipfel S, Herzog W. Gesundheitsfragebogen für patienten (PHQ-D). 2nd ed. Karlsruhe: Pfizer: Manual und Testunterlagen; 2002.
[22] Alexopoulos GS, Raue PJ, Kiosses DN, Mackin RS, Kanellopoulos D, McCulloch C, et al. Problem-solving therapy and supportive therapy in older adults with major depression and executive dysfunction. Arch Gen Psychiatry 2011;68 (1):33–41.
[23] D'Zurilla TJ, Goldfried MR. Problem solving and behavior modification. J Abnorm Psychol 1971;78(1):107–26.
[24] Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. J Gen Intern Med. 2001;16(9):606–13.
[25] Spitzer RL, Kroenke K, Williams JB. Validation and utility of a self-report version of PRIME-MD: the PHQ primary care study. Primary care evaluation of mental disorders. Patient health questionnaire. JAMA 1999;282(18):1737–44.
[26] Gräfe K, Zipfel S, Herzog W, Löwe B. Screening psychischer störungen mit dem "Gesundheitsfragebogen für patienten (PHQ-D)". Diagnostica 2004;50 (4):171–81.
[27] The EuroQol Group. EuroQol - a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.
[28] König H-H, Heider D, Lehnert T, Riedel-Heller SG, Angermeyer MC, Matschinger H, et al. The ESEMeD/MHEDEA investigators. Health status of the advanced elderly in six European countries: results from a representative survey using EQ-5D and SF-12. Health Qual Life Outcomes 2010;8:143.
[29] van Leeuwen KM, Bosmans JE, Jansen AP, Hoogendijk EO, van Tulder MW, van der Horst HE, et al. Comparing measurement properties of the EQ-5D-3L, ICECAP-O, and ASCOT in frail older adults. Value Health 2015;18(1):35–43.
[30] Dolan P. Modeling valuations for EuroQol health states. Med Care 1997;35 (11):1095–108.
[31] Seidl H, Bowles D, Bock JO, Brettschneider C, Greiner W, König H-H, et al. FIMA-fragebogen zur erhebung von gesundheitsleistungen im alter: entwicklung und pilotstudie. Gesundheitswesen 2015;77(1):46–52.
[32] Kroenke K, Spitzer RL, Williams JBW, Löwe B. The patient health questionnaire somatic, anxiety, and depressive symptom scales: a systematic review. Gen Hosp Psychiatry 2010;32(4):345–59.
[33] Rigler SK, Studenski S, Wallace D, Reker DM, Duncan PW. Co-morbidity adjustment for functional outcomes in community-dwelling older adults. Clin Rehabil 2002;16(4):420–8.
[34] Vannoy SD, Arean P, Unützer J. Advantages of using estimated depression-free days for evaluating treatment efficacy. Psychiatr Serv 2010;61(2):160–3.
[35] Mallick R, Chen J, Entsuah AR, Schatzberg AF. Depression-free days as a summary measure of the temporal pattern of response and remission in the treatment of major depression: a comparison of venlafaxine, selective serotonin reuptake inhibitors, and placebo. J Clin Psychiatry 2003;64 (3):321–30.
[36] Lave JR, Frank RG, Schulberg HC, Kamlet MS. Cost-effectiveness of treatments for major depression in primary care practice. Arch Gen Psychiatry 1998;55 (7):645–51.
[37] Bock JO, Brettschneider C, Seidl H, Bowles D, Holle R, Greiner W, et al. Ermittlung standardisierter Bewertungssätze aus gesellschaftlicher Perspektive für die gesundheitsökonomische evaluation. Gesundheitswesen 2015;77(1):53–61.
[38] Statistisches Bundesamt. Preise. Verbraucherpreisinidzes für Deutschland. Lange reihen ab 1948 April 2015. Wiesbaden: Statistisches Bundesamt; 2015.
[39] The World Bank. World development indicators: size of the economy. Washington, DC: The World Bank; 2013.
[40] Die Rote Liste Service GmbH. ROTE LISTE 2014: Arzneimittelverzeichnis für Deutschland (einschließlich EU-Zulassungen und bestimmter Medizinprodukte). Frankfurt am Main: Die Rote Liste Service GmbH; 2014.
[41] Statistisches Bundesamt. Verdienste und Arbeitskosten. Verdienststrukturen 2010. Wiesbaden: Statistisches Bundesamt; 2013.
[42] Statistisches Bundesamt. Verdienste und Arbeitskosten. Arbeitskosten im Produzierenden Gewerbe und im Dienstleistungsbereich - Ergebnisse für Deutschland- 2012. Wiesbaden: Statistisches Bundesamt; 2015.
[43] StataCorp. Stata multiple impuatation reference manual. Release 13. College Station, TX: StataCorp LP; 2013.

https://doi.org/10.1016/j.eurpsy.2018.12.007 Published online by Cambridge University Press

[44] Drummond MF, Sculpher MJ, Torrance GW, O'Brien BJ, Stoddart GL. Methods for the economic evaluation of health care programmes. 3rd ed. Oxford, New York: Oxford University Press; 2005.
[45] Zellner A. An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. J Am Stat Assoc 1962;57(298) 348-&.
[46] Tambour M, Zethraeus N, Johannesson M. A note on confidence intervals in cost-effectiveness analysis. Int J Technol Assess Health Care 1998;14 (3):467–71.
[47] Lothgren M, Zethraeus N. Definition, interpretation and calculation of cost-effectiveness acceptability curves. Health Econ 2000;9(7):623–30.
[48] Brazier J, Connell J, Papaioannou D, Mukuria C, Mulhern B, Peasgood T, et al. A systematic review, psychometric analysis and qualitative assessment of generic preference-based measures of health in mental health populations and the estimation of mapping functions from widely used specific measures. Health Technol Assess 2014;18(34):1–188.
[49] Grochtdreis T, Brettschneider C, Hajek A, Schierz K, Hoyer J, König H-H. Mapping the beck depression inventory to the EQ-5D-3L in patients with depressive disorders. J Ment Health Policy Econ 2016;19(2):79–89.
[50] Hoch JS, Briggs AH, Willan AR. Something old, something new, something borrowed, something blue: a framework for the marriage of health econometrics and cost-effectiveness analysis. Health Econ 2002;11(5):415–30.
[51] Unützer J, Katon WJ, Russo J, Simon G, Von Korff M, Lin E, et al. Willingness to pay for depression treatment in primary care. Psychiatr Serv 2003;54(3):340–5.
[52] Rubin DB. Multiple imputation for nonresponse in surveys. New York: John Wiley & Sons; 1987.
[53] Little RJ, Rubin DB. Causal effects in clinical and epidemiological studies via potential outcomes: concepts and analytical approaches. Annu Rev Public Health 2000;21:121–45.
[54] Huijbregts KM, de Jong FJ, van Marwijk HW, Beekman AT, Ader HJ, Hakkaart-van Roijen L, et al. A target-driven collaborative care model for Major depressive disorder is effective in primary care in the Netherlands. A randomized clinical trial from the depression initiative. J Affect Disord 2013;146(3):328–37.
[55] Richards DA, Hill JJ, Gask L, Lovell K, Chew-Graham C, Bower P, et al. Clinical effectiveness of collaborative care for depression in UK primary care (CADET): cluster randomised controlled trial. BMJ 2013;347:f4913.

https://doi.org/10.1016/j.eurpsy.2018.12.007 Published online by Cambridge University Press