Psychological Medicine (2017), **47**, 1825–1835. © Cambridge University Press 2017
doi:10.1017/S0033291717000289

ORIGINAL ARTICLE

Cost-effectiveness of computerized cognitive–behavioural therapy for the treatment of depression in primary care: findings from the Randomised Evaluation of the Effectiveness and Acceptability of Computerised Therapy (REEACT) trial

A. Duarte[1]*, S. Walker[1], E. Littlewood[2], S. Brabyn[2], C. Hewitt[2], S. Gilbody[2] and S. Palmer[1] on behalf of the REEACT team

[1] *Centre for Health Economics, University of York, Heslington, York YO10 5DD, UK*
[2] *Department of Health Sciences, University of York, Heslington, York YO10 5DD, UK*

Background. Computerized cognitive–behavioural therapy (cCBT) forms a core component of stepped psychological care for depression. Existing evidence for cCBT has been informed by developer-led trials. This is the first study based on a large independent pragmatic trial to assess the cost-effectiveness of cCBT as an adjunct to usual general practitioner (GP) care compared with usual GP care alone and to establish the differential cost-effectiveness of a free-to-use cCBT programme (MoodGYM) in comparison with a commercial programme (Beating the Blues) in primary care.

Method. Costs were estimated from a healthcare perspective and outcomes measured using quality-adjusted life years (QALYs) over 2 years. The incremental cost-effectiveness of each cCBT programme was compared with usual GP care. Uncertainty was estimated using probabilistic sensitivity analysis and scenario analyses were performed to assess the robustness of results.

Results. Neither cCBT programme was found to be cost-effective compared with usual GP care alone. At a £20 000 per QALY threshold, usual GP care alone had the highest probability of being cost-effective (0.55) followed by MoodGYM (0.42) and Beating the Blues (0.04). Usual GP care alone was also the cost-effective intervention in the majority of scenario analyses. However, the magnitude of the differences in costs and QALYs between all groups appeared minor (and non-significant).

Conclusions. Technically supported cCBT programmes do not appear any more cost-effective than usual GP care alone. No cost-effective advantage of the commercially developed cCBT programme was evident compared with the free-to-use cCBT programme. Current UK practice recommendations for cCBT may need to be reconsidered in the light of the results.

Received 31 March 2016; Revised 19 January 2017; Accepted 23 January 2017; First published online 23 February 2017

Key words: Computerized cognitive–behavioural therapy, depression, primary care.

Introduction

Depression is a highly prevalent condition that makes a considerable impact on patients' health-related quality of life (HRQoL) (Moussavi *et al.* 2007). It is one of the most common reasons for consulting with a general practitioner (GP) and leads to the expenditure of large amounts of healthcare resources (Üstün *et al.* 2004). The burden of depression is further increased, as incomplete recovery and relapse are common, with a risk of relapse as high as 50% following the first episode rising to 70% for those who experience a second episode (Kupfer *et al.* 1996).

Current clinical guidelines in the UK recommend a 'stepped-care approach' to depression management depending on severity, response to treatment and patient preference. Psychosocial interventions, such as cognitive–behavioural therapy (CBT), behavioural activation and problem solving, in combination with other treatments are recommended at different levels of intensity for: step 1, all forms of depression (suspected or known); step 2, persistent subthreshold depressive symptoms and mild to moderate depression; step 3, severe depression or lower-severity

* Address for correspondence: A. Duarte, Centre for Health Economics, University of York, Heslington, York YO10 5DD, UK.
(Email: ana.duarte@york.ac.uk)

depression not responsive to step 2 treatment; and step 4, severe and complex depression. Patients are offered the least intrusive, most effective treatment according to their presentation of depression, and move up the steps upon treatment failure or if they decline the offered intervention. This constitutes the standard of care in the UK, and it is accessed through primary care or self-referral, with the GP as gatekeeper to more specialized levels of care (National Institute for Health and Clinical Excellence, 2009). Amongst these psychological interventions, CBT has been identified as a leading evidence-supported form of brief psychological therapy for people with depression (Roth & Fonagy, 2005; National Institute for Health and Clinical Excellence, 2009). However, the scarcity of therapists leads to under-provision of face-to-face CBT, (Bower & Gilbody, 2005) and computer-delivered CBT (cCBT) can constitute an alternative (Kaltenthaler *et al.* 2006). In the UK, cCBT is currently recommended as a low-intensity intervention at step 2, i.e. for persistent subthreshold depressive symptoms and mild to moderate depression (National Institute for Health and Clinical Excellence, 2009).

cCBT is currently part of step 2 in the National Institute for Health and Care Excellence's (NICE) 'stepped approach', as a form of low-intensity psychosocial therapy for the treatment of depression in primary care (National Institute for Health and Clinical Excellence, 2009). This recommendation was largely informed by clinical and cost-effectiveness data from developer-led trials (Christensen *et al.* 2004; McCrone *et al.* 2004; Proudfoot *et al.* 2004) and pertained to two cCBT programmes, commercial Beating the Blues and free-to-use MoodGYM. Furthermore, existing effectiveness evidence suggests that cCBT (commercial and free to use) is comparable with therapist-delivered cCBT (Kaltenthaler *et al.* 2006; Spek *et al.* 2007; Andersson & Cuijpers, 2009).

Concerns about the generalizability and external validity of the data used to inform these clinical guidelines have led to recommendations for further studies which: (1) recruit participants in primary care settings (rather than academic centres or secondary care); and (2) follow-up patients beyond 1 year (Andersson & Cuijpers, 2009).

The Randomised Evaluation of the Effectiveness and Acceptability of Computerised Therapy (REEACT) trial was conducted in response to the need for independent clinical and cost-effectiveness evaluation of cCBT in a primary care setting, and a longer-term follow-up period. The trial methodology and the clinical results have been previously reported (Gilbody *et al.* 2015; Littlewood *et al.* 2015). An integral part of the design of this study was the inclusion of an economic study to assess the cost-effectiveness of cCBT when added to usual GP care (as defined by NICE guidance) (National Institute for Health and Clinical Excellence, 2009), compared with usual GP care alone. Importantly, this was a large trial ($n = 691$) with statistical power exceeding those of prior studies to detect clinically significant treatment effects, and including patient resource use as well as HRQoL assessment using two generic preference-based instruments recognized as suitable to inform economic evaluation, the three-level EuroQol five dimensions questionnaire (EQ-5D-3L) (The EuroQol Group, 1990; Brooks, 1996) and the Six-Dimension Short-Form (SF-6D) (Brazier *et al.* 1998, 2002). This paper reports the results of the cost-effectiveness analysis based on the REEACT trial and examines the incremental benefits of adding cCBT to usual GP care from an economic perspective.

Method

The primary objective of the economic analysis was to assess the cost-effectiveness of cCBT as an adjunct to usual GP care compared with usual GP care alone and to establish the differential cost-effectiveness of a free-to-use cCBT programme (MoodGYM) in comparison with a commercial pay-to-use cCBT programme (Beating the Blues). The economic analysis was conducted prospectively alongside a randomized controlled trial (RCT) in a primary care setting (REEACT). The methodology of the trial has been described in detail elsewhere and is summarized in brief below (Gilbody *et al.* 2015; Littlewood *et al.* 2015).

The trial is registered as Current Controlled Trials ISRCTN91947481.

Study design and participants

The REEACT trial was a pragmatic, multicentre, open, three-armed, parallel RCT conducted in nine study sites across England in a primary care setting. The trial was designed to test the effectiveness of technically supported cCBT when added to usual GP care, and also to test the non-inferiority of free-to-use cCBT compared with commercially developed cCBT. A total of 691 adults presenting with depression according to a self-report questionnaire [score of $\geqslant 10$ on the Patient Health Questionnaire (PHQ-9) depression severity instrument; Kroenke *et al.* 2001] who were not in receipt of cCBT or specialist psychological therapy at the time of recruitment were included in the trial. Participants were excluded if they were actively suicidal, suffering from psychotic symptoms (ascertained by GP), depressed in the postnatal period, had suffered bereavement within the last year, had a primary diagnosis of alcohol or drug abuse or were not

https://doi.org/10.1017/S0033291717000289 Published online by Cambridge University Press

able to read and write in English. Participants were followed up for 24 months, and data were collected from participants at baseline (prior to randomization), and at 4, 12 and 24 months post-randomization.

Interventions

A total of 691 participants were randomized to receive either usual GP care ($n = 239$) or usual care from their GP plus one of two interventions: (i) Beating the Blues ($n = 210$) or (2) MoodGYM ($n = 242$). Both programmes had previously been recommended in clinical guidelines (National Institute for Health and Clinical Excellence, 2009) and had been shown to be clinically and cost-effective based on developer-led trials (Christensen *et al.* 2004; McCrone *et al.* 2004; Proudfoot *et al.* 2004). All participants randomized to the cCBT programmes continued to receive the standard care they would have received from their GP if the trial had not been in place. No restrictions were imposed on usual care, with treatment being provided at the GP discretion. The cCBT programmes were supported by weekly telephone calls delivered by trained technicians, so as to provide technical support on the cCBT programmes and to encourage participants to engage with the programmes. The support provided replicated or exceeded the support offered in routine National Health Service (NHS) primary care psychological therapy services. In view of the pragmatic nature of the trial, treatments were not constrained. Following randomization, participants in usual care and cCBT arms were free to consult with their GP and were able to access the full range of additional forms of psychological therapy or drug treatment that would otherwise be available to people with depression in primary care.

Outcomes

The primary clinical outcome in the REEACT trial was self-reported symptoms of depression at 4 months assessed using a validated depression severity instrument (PHQ-9). Secondary outcomes included self-reported symptoms of depression at 12 and 24 months. Full details and results of the primary and secondary outcomes have been reported in detail elsewhere (Gilbody *et al.* 2015; Littlewood *et al.* 2015). In summary, participants offered commercial or free-to-use cCBT experienced no additional improvement in depression compared with usual GP care at 4 months [Beating the Blues *v.* usual GP care odds ratio (OR) 1.19, 95% confidence interval (CI) 0.75–1.88; MoodGYM *v.* usual GP care OR 0.98, 95% CI 0.62–1.56]. In a repeated-measures analysis across all time points there was no statistical evidence of an overall difference between Beating the Blues or MoodGYM compared with usual GP care (OR 0.99, 95% CI 0.57–1.70 and OR 0.68, 95% CI 0.42–1.10, respectively).

A potential limitation of using self-reported symptoms of depression in a cost-effectiveness analysis is that this precludes comparison of the cost-effectiveness of cCBT with other interventions seeking NHS funding. The use of a single, generic measure of health benefit enables diverse healthcare interventions to be compared, thus enabling broader questions of efficiency to be addressed. Consequently, the main outcome for the cost-effectiveness analysis was the quality-adjusted life year (QALY) assessed using two standardized generic and preference-based measures: the EQ-5D-3L (Brooks, 1996; The EuroQol Group, 1990) and the SF-6D (derived from the Short-Form-36; SF-36; Brazier *et al.* 1998, 2002). These were completed at baseline and at 4, 12 and 24 months post-randomization. The scores at each time point were used to estimate QALYs using the area under the curve method, which multiplies HRQoL weights by time (Matthews *et al.* 1990). QALYs accrued from 12 to 24 months were discounted at a 3.5% discount rate, in line with current UK guidance (National Institute for Health and Care Excellence, 2013).

In the base-case analysis we estimated QALYs based on the EQ-5D-3L as this forms part of the reference case for cost-effectiveness studies submitted to NICE (National Institute for Health and Care Excellence, 2013). This EQ-5D-3L asks participants to rate the severity of their problems (no problem, moderate problems or severe problems) in five health domains: mobility, self-care, usual activities, pain/discomfort and anxiety/depression. These ratings define health states which have been assigned preference weights using preferences measured in a representative sample of the UK population (The EuroQol Group, 1990; Dolan *et al.* 1995). As part of a separate scenario analysis, QALYs were also estimated using the SF-6D preference scores generated from participants' response to the SF-36v2 (Brazier *et al.* 2002).

Resource use and costs

Healthcare resource use data were obtained via objective data collection from GP medical records, and collected from 2 months pre-randomization to 24-month post-randomization follow-up. The data were obtained across three time-frames: (1) from 2 months pre-randomization to the date of randomization ('baseline'); (2) from the date of randomization to the 12-month follow-up ('year 1'); and (3) from the 12-month follow-up to the 24-month follow-up ('year 2'). Data were collected on the following healthcare resource use items: number of primary care consultations (GP and nurse); depression-related prescribed

https://doi.org/10.1017/S0033291717000289 Published online by Cambridge University Press

medication (antidepressants, antipsychotics, mood stabilizers, sedatives and anxiolytics); referrals to other community mental health services and number of sessions (counsellors, community mental health teams, improving access to psychological therapies, psychologists, psychiatrists); in-patient hospital admissions and length of stay; out-patient hospital appointments; number of emergency contacts, including accident and emergency attendances and contacts made with out-of-hours services. The number and duration of telephone support calls by treatment arm were recorded as part of the study by three telephone support workers. Researchers who conducted data collection and staff providing telephone support were not blind to treatment allocation.

Healthcare costs were estimated by multiplying the resource use by the appropriate unit cost, using routinely published UK unit cost estimates (pounds sterling at 2011–2012 prices) (Curtis, 2012; Department of Health, 2012; Joint Formulary Committee, 2013). The costs associated with the provision of cCBT include the licence fee (applicable only to Beating the Blues) and the cost of telephone support (Supplementary Table S2; online Supplementary material). All costs related to the provision of cCBT were assumed to be incurred in the first year of follow-up (year 1). Costs accrued from 12 to 24 months were also discounted at a 3.5% discount rate (National Institute for Health and Care Excellence, 2013).

Analysis

The cost-effectiveness analysis was conducted from a healthcare provider perspective on an intention-to-treat basis and with a time horizon of 24 months. We estimated the mean healthcare costs incurred and QALYs accrued in each treatment group using regression analyses controlling for pre-specified covariates (age, sex, anxiety level at baseline, depression severity at baseline, and depression duration at baseline). For QALYs, baseline EQ-5D was also controlled for (Manca *et al.* 2005), and similarly the costs regression was controlled for baseline costs. To account for missing data, we used multiple imputation methods with chained equations (Royston, 2004) and predictive mean matching over 10 imputations to estimate cost aggregated by resource use category (see above) and EQ-5D-3L and SF-6D data items when these were missing. EQ-5D-3L and SF-6D scores were imputed at every follow-up time point (baseline, 4, 12 and 24 months) whilst costs by category were imputed for the same time intervals as the resource use data collection (2 months prior to randomization, from randomization to 12 months, and from 12 to 24 months). The independent variables specified in the imputation were: baseline EQ-5D-3L score, baseline SF-6D score, baseline costs, age, sex, anxiety level at baseline, depression severity at baseline, and depression duration at baseline.

Mean differences in total costs and QALYs were estimated for each cCBT programme *v*. usual GP care using regression analysis to control for age, sex, anxiety level, depression severity, and depression duration at baseline (covariates used in the clinical effectiveness analyses), as well as baseline costs for total costs and baseline EQ-5D-3L score for QALYs (Manca *et al.* 2005). The regression model selected for all cost analysis was a generalized linear model (GLM) with an identity link function and a γ distribution for error terms (Barber & Thompson, 2004). This type of model was preferred to an ordinary least squares (OLS) regression, as cost data tend to be heavily skewed and follow a non-normal distribution and are thus likely to violate the underlying assumptions of OLS. For mean differences in QALYs, OLS regression was used.

In the base-case analysis we calculated the additional cost per QALY gained (incremental cost-effectiveness ratio; ICER) of each cCBT intervention compared with usual GP care based on mean QALYs generated from EQ-5D-3L scores and mean total costs of healthcare utilization. The ICER was compared with the lower bound of the cost-effectiveness threshold range of £20 000 to £30 000 per additional QALY (threshold range adopted by NICE) (National Institute for Health and Care Excellence, 2013). Probabilistic sensitivity analysis was performed to estimate decision uncertainty based on all three treatment options; that is, the probability that the joint uncertainty in costs and QALYs would lead to each intervention being cost-effective at a given cost-effectiveness threshold, and presented these probabilities in cost-effectiveness acceptability curves (CEACs) (Fenwick *et al.* 2001).

In order to plot the CEAC, the variance–covariance matrices from the costs and QALYs regressions were extracted and the corresponding Cholesky decompositions used to obtain correlated draws from a multivariate normal distribution (Briggs *et al.* 2006).

Three scenario analyses were performed to assess the robustness of the findings to alternative assumptions regarding source of HRQoL, costs and missing data. Scenario 1 used alternative QALY estimates generated from SF-6D scores. In scenario 2, only costs related to depression were included in the cost analysis; total depression-related costs included depression-related costs of GP and nurse visits, other mental health community services attendances and depression-related medication costs. In scenario 3, only participants with complete data were included.

All analyses were conducted using STATA/SE version 12.0 (Stata Statistical Software: release 12; StataCorp LP, USA) and Microsoft Excel 2010.

https://doi.org/10.1017/S0033291717000289 Published online by Cambridge University Press

Table 1. *EQ-5D summary scores and QALYs estimated on the multiple imputed data sets (adapted from the Health Technology Assessment report) (Littlewood* et al. *2015)*

	Beating the Blues		Usual GP care		MoodGYM	
	Mean (S.E.)	*n* (%)[a]	Mean (S.E.)	*n* (%)[a]	Mean (S.E.)	*n* (%)[a]
Baseline	0.6162 (0.0173)	210 (100.0)	0.6256 (0.0174)	239 (100.0)	0.6111 (0.0173)	242 (100.0)
4 months	0.6854 (0.0204)	157 (74.7)	0.7275 (0.0184)	167 (69.9)	0.6851 (0.0211)	170 (70.2)
12 months	0.6914 (0.0225)	144 (68.6)	0.7081 (0.0210)	156 (65.3)	0.7186 (0.0203)	159 (65.7)
24 months	0.6677 (0.0234)	129 (61.4)	0.7093 (0.0216)	143 (59.8)	0.6734 (0.0250)	144 (59.5)
QALYs[b]	1.3325 (0.0337)	N.A.	1.3888 (0.0328)	N.A.	1.3564 (0.0330)	N.A.

EQ-5D, EuroQol five dimensions questionnaire; QALYs, quality-adjusted life years; GP, general practitioner; S.E., standard error; N.A., not applicable.

[a] Number of participants with a reported EQ-5D score.

[b] QALYs in year 2 were discounted at a 3.5% rate.

Results

Sample characteristics

Participant characteristics at baseline were similar across the three groups, in terms of age, sex, severity of depression, duration of depression, anti-depressant use and educational attainment. The majority of participants were female (87%) and the mean age was 39.9 years. The median PHQ-9 score was 17 across the groups, indicating moderate depression severity (Kroenke *et al.* 2001). Further details of participants can be found in the online Supplementary material and elsewhere (Littlewood *et al.* 2015).

Outcomes

Health outcomes in terms of EQ-5D-3L scores at each time point and QALYs accrued over the trial period based on the imputed data are shown in Table 1. Corresponding results for the SF-6D are reported in the online Supplementary material. Unadjusted mean estimates of QALYs over 24 months based on the EQ-5D-3L were 1.3325 (S.E. = 0.0337) for Beating the Blues, 1.3888 (S.E. = 0.0328) for usual GP care, and 1.3564 (S.E. = 0.0330) for MoodGYM.

Resource use and costs

Descriptive statistics of healthcare resource use over the 24 months follow-up period based on the available case dataset and also the unit costs associated with each category of resource use are shown in Table 2. Costs associated with the delivery of cCBT programmes are reported in the online Supplementary material. Overall, the proportion of available GP records from which resource use data were extracted was of similar magnitude for Beating the Blues (82.4%), usual GP care alone (84.5%) and MoodGYM (84.7%). In general, differences between treatment groups in resource use appeared small, although resource use estimates across participants were considerably variable with large standard deviations.

Table 3 reports the mean costs for each of the major types of service. Primary care services represented the largest share of healthcare expenditure for all treatment groups, comprising over 50% of total costs for all groups. The second largest category of costs was hospital services which varied from 25% to 35% across the groups. Mean total unadjusted costs for the 24-month period were £1186 (S.E. = £79) for Beating the Blues, £1121 (S.E. = £61) for usual GP care alone, and £1098 (S.E. = £134) for MoodGYM.

Cost-effectiveness analysis

Base-case analysis

Mean differences in total costs and QALYs for each cCBT programme *v.* usual GP care alone with adjustment for covariates are reported in Table 4. In general, differences in costs and QALYs between both cCBT groups and usual GP care were small with wide CIs, and were not statistically significant at the 5% significance level.

The base-case results suggest that neither Beating the Blues nor MoodGYM plus usual GP care appeared cost-effective compared with usual GP care alone. Based on a comparison of the mean differences in total costs and QALYs, Beating the Blues plus usual GP care appears dominated by usual GP care alone, with higher mean costs and lower QALYs. MoodGYM resulted in both lower mean costs and QALYs compared with usual GP care. Therefore, the ICER estimated falls within the south-west quadrant of the cost-effectiveness plane. In this quadrant, the interpretation of the ICER refers to the difference in costs and QALYs between the higher-cost

https://doi.org/10.1017/S0033291717000289 Published online by Cambridge University Press

Table 2. *Resource use from randomization to 24 months of follow-up (adapted from the main report) (Littlewood* et al. *2015)*

	Beating the Blues ($n = 173$)[a]			Usual GP care ($n = 202$)[a]			MoodGYM ($n = 205$)[a]				
Service	Mean (S.D.)	Median	Used by, %[b]	Mean (S.D.)	Median	Used by, %[b]	Mean (S.D.)	Median	Used by, %[b]	Unit	Unit cost or range, £
Primary care											
GP	13.91 (9.23)	12.00	98.15	12.03 (8.15)	11.00	96.34	12.24 (8.55)	10.00	98.46	Contact	26–110
Nurse	3.56 (4.45)	2.00	75.46	3.76 (5.14)	2.00	74.87	3.88 (5.85)	2.00	72.96	Contact	2.10–32.54
Out of hours	0.24 (0.70)	0.00	16.56	0.14 (0.49)	0.00	10.47	0.12 (0.41)	0.00	9.18	Contact	61.14
Hospital											
In-patient	0.36 (1.55)	0.00	9.62	0.36 (1.36)	0.00	10.05	0.66 (5.29)	0.00	9.28	LoS	206.89–430.14
Out-patient	2.10 (3.68)	1.00	12.43	1.54 (1.96)	1.00	8.47	1.84 (2.99)	1.00	10.36	Attendance	14–459
Other community services											
IAPT	0.31 (1.35)	0.00	6.79	0.45 (1.27)	0.00	14.66	0.38 (1.53)	0.00	8.16	Visit	70.35
CMHT	0.21 (1.62)	0.00	5.52	0.30 (1.36)	0.00	10.47	0.11 (0.46)	0.00	6.12	Visit	44.84
Counsellor	0.33 (1.51)	0.00	8.02	0.10 (0.56)	0.00	4.19	0.19 (1.02)	0.00	4.59	Visit	59
Psychiatric	0.24 (2.01)	0.00	4.91	0.17 (0.89)	0.00	4.19	0.12 (0.90)	0.00	2.55	Visit	174.64
Psychological	0.51 (3.34)	0.00	4.29	0.56 (2.83)	0.00	6.28	0.03 (0.24)	0.00	1.53	Visit	70.08
Medication											
Depression related	N.A.	N.A.	83.13	N.A.	N.A.	85.64	N.A.	N.A.	82.38	Item	Various

GP, General practitioner, S.D., standard deviation; LoS, length of stay; IAPT, improving access to psychological therapies; CMHT, community mental health team; N.A, not applicable.

[a] Number of participants for whom any resource use data in GP records were available.

[b] Percentage of use within number of participants with available data in each resource use category.

Table 3. *Summary of costs during trial follow-up (adapted from the Health Technology Assessment report) (Littlewood* et al. *2015)*

	Beating the Blues		Usual GP care		MoodGYM	
Costs, £[a]	Mean (S.E.)	%	Mean (S.E.)	%	Mean (S.E.)	%
Primary care	628.58 (32.32)	52.98	556.86 (28.61)	54.55	563.58 (28.45)	51.30
Hospital services	302.60 (43.20)	25.51	277.36 (35.54)	27.17	383.98 (116.07)	34.95
Other community services	129.53 (31.28)	10.92	111.48 (20.86)	10.92	72.10 (17.27)	6.56
Medication	75.28 (11.70)	6.35	75.05 (12.51)	7.35	78.22 (12.86)	7.12
cCBT	50.43 (0.00)	4.25	0.00 (0.00)	0.00	0.66 (0.00)	0.06
Total costs	1186.43 (78.79)	100.00	1020.76 (60.56)	100.00	1098.53 (134.04)	100.00

GP, General practitioner, S.E., standard error; cCBT, computerized cognitive–behavioural therapy.
[a] Costs in year 2 were discounted at a 3.5% rate.

Table 4. *Results of incremental cost-effectiveness analysis of MoodGYM and Beating the Blues compared with usual GP care over 24 months (adapted from the Health Technology Assessment report; Littlewood* et al. *2015)*

Analysis	Δ Costs, £ (95% CI)	Δ QALYs (95% CI)	ICER, £ per QALY[a]	Probability of CE at £20 000/QALY
Base-case				
Usual GP care	–	–	–	0.545
Beating the Blues	104.24 (−66.74 to 275.26)	−0.0435 (−0.1167 to 0.0297)	Dominated	0.038
MoodGYM	−106.07 (−261.65 to 49.52)	−0.0153 (−0.0919 to 0.0613)	6933[b]	0.417
Scenario 1 – SF-6D as HRQoL source				
Usual GP care	–	–	–	0.237
Beating the Blues	104.24 (−66.74 to 275.26)	−0.0277 (−0.0672 to 0.0118)	Dominated	0.007
MoodGYM	−106.07 (−261.65 to 49.52)	0.0058 (−0.0294 to 0.0409)	Dominant	0.756
Scenario 2 – depression costs only				
Usual GP care	–	–	–	0.581
Beating the Blues	52.55 (−45.48 to 150.58)	−0.0277 (−0.0672 to 0.0118)	Dominated	0.051
MoodGYM	−57.15 (−142.98 to 28.67)	−0.0153 (−0.0919 to 0.0613)	3735[b]	0.368
Scenario 3 – complete case analysis				
Usual GP care	–	–	–	0.601
Beating the Blues	−23.62 (−247.75 to 200.51)	−0.0485 (−0.1488 to 0.05178)	487[b]	0.109
MoodGYM	−176.54 (−386.44 to 33.37)	−0.0295 (−0.1263 to 0.0674)	5984[b]	0.290

GP, General practitioner; CI, confidence interval; QALYs, quality-adjusted life years; ICER, incremental cost-effectiveness ratio; CE, cost-effectiveness; SF-6D, Six-Dimension Short-Form; HRQoL, health-related quality of life.
[a] Compared with usual GP care.
[b] ICER on the south-west quadrant of the CE plane (ICER refers to cost-effectiveness of usual GP care alone *v.* intervention).

intervention (usual GP care) and the lower-cost intervention (in this case, MoodGYM). Consequently, the ICER of £6933 per additional QALY represents the ICER of usual GP care alone *v.* MoodGYM plus usual GP care (Table 4). Since this falls below the £20 000 per QALY threshold, usual GP care is considered more cost-effective than MoodGYM.

Table 4 also reports the probability of cost-effectiveness for each treatment. At a £20 000 per QALY threshold, usual GP care appears the treatment most likely to be cost-effective followed by MoodGYM plus GP care then Beating the Blues plus GP care (with a 0.545, 0.417 and 0.038 probability of cost-effectiveness, respectively). The probability of each

https://doi.org/10.1017/S0033291717000289 Published online by Cambridge University Press

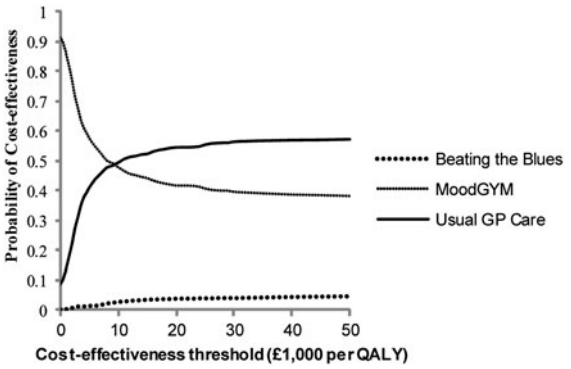

Fig. 1. Cost-effectiveness acceptability curves for the three interventions (adapted from the Health Technology Assessment report; Littlewood *et al.* 2015). GP, General practitioner; QALY, quality-adjusted life year.

intervention being cost-effective at a range of cost-effectiveness thresholds is shown in Fig. 1.

Scenario analyses

Table 4 presents the results of the three scenario analyses. Using SF-6D values for HRQoL and QALYs (scenario 1), Beating the Blues plus GP care remained dominated by usual GP care alone, which was consistent with the base-case. In contrast, MoodGYM plus GP care had a positive, albeit small, QALY increment (0.0058) compared with usual GP care alone, whilst remaining cost saving, therefore dominating usual GP care alone. Thus MoodGYM appears to be cost-effective, resulting in lower mean costs and higher mean QALYs compared with usual GP care. Nevertheless, the estimates were not statistically significant at the 5% significance level for either comparison of cCBT against usual GP care. At a £20 000 per QALY threshold, MoodGYM had a 0.756 probability of being the optimal intervention in terms of cost-effectiveness.

Where only depression-related treatment costs were included (scenario 2), the incremental costs were consistent with the main analysis for both cCBT programmes, although the magnitude of the differences between the cCBT groups and usual GP care was reduced. Usual GP care was also the cost-effective intervention in the complete data analysis (scenario 3); however, Beating the Blues was not dominated in this scenario. Full incremental results for the scenario analyses are shown alongside the base-case in Table 4.

Discussion

The study suggests that neither MoodGYM, nor Beating the Blues appears cost-effective when added to usual GP care and compared with usual GP care alone for the management of depression in primary care. These findings were robust to alternative assumptions on costs and missing data with the exception of the choice of the HRQoL instrument. When the SF-6D was used instead of the EQ-5D, MoodGYM appeared to dominate usual GP care alone (lower mean costs and higher QALYs) and was the intervention most likely to be cost-effective at a £20 000 per QALY threshold. However, differences in the mean cost and QALY estimates were not statistically significant using either the EQ-5D-3L or SF-6D for either comparison of cCBT against usual GP care. A consistent finding across all scenarios was that the commercially developed programme (Beating the Blues) conferred no additional health economic benefit compared with the free-to-use programme (MoodGYM).

It is important to consider why the results are sensitive to the choice of the HRQoL measurement instrument, as NICE also accepts the use of the SF-6D when EQ-5D measured utilities are not available. Nevertheless, it has been demonstrated that, despite the convergence of measurements by the EQ-5D-3L and SF-6D, the two instruments are not interchangeable (Brazier *et al.* 2004). Whilst the results appear sensitive to the choice of whether the EQ-5D-3L or SF-6D is used to estimate QALYs, the differences between all three groups were relatively minor both in terms of costs and QALYs. Hence minor differences in the assumptions can lead to different cost-effectiveness interpretations due to relatively small impacts on the mean incremental estimates of costs and QALYs, and results should be interpreted cautiously.

The lack of a statistically significant improvement in terms of QALYs associated with the addition of cCBT to usual GP care may be because neither of the generic quality-of-life instruments (EQ-5D-3L and SF-6D) was sufficiently sensitive to changes in the quality of life in this patient group. However, it appears more likely that the use of cCBT has a negligible impact on patient quality of life in comparison with usual GP care alone and appears consistent with the findings reported for the primary clinical outcome reported in the main trial paper where there were no discernible clinical benefits of cCBT in terms of depression outcomes (Gilbody *et al.* 2015).

Our findings are in contrast to those of previous studies that identified cCBT interventions as cost-effective (McCrone *et al.* 2004; Kaltenthaler *et al.* 2006; National Institute for Health and Clinical Excellence, 2009; Gerhards *et al.* 2010; Hollinghurst *et al.* 2010; Warmerdam *et al.* 2010). There are important differences in these other economic evaluations that may explain the discrepancy in results with the REEACT study, such as shorter durations of patient follow-up in previous economic evaluations and trials that

https://doi.org/10.1017/S0033291717000289 Published online by Cambridge University Press

informed them (Christensen *et al.* 2004; McCrone *et al.* 2004; Gerhards *et al.* 2010; Hollinghurst *et al.* 2010; Warmerdam *et al.* 2010), smaller sample size in earlier studies (McCrone *et al.* 2004; Gerhards *et al.* 2010; Hollinghurst *et al.* 2010; Warmerdam *et al.* 2010), estimation of QALYs by mapping from a depression-specific measure (Beck's Depression Inventory; McCrone *et al.* 2004; Hollinghurst *et al.* 2010), intervention delivered online by a therapist (Hollinghurst *et al.* 2010) and use of a different analytic perspective (societal) which included non-healthcare costs in the analysis (Gerhards *et al.* 2010; Hollinghurst *et al.* 2010; Warmerdam *et al.* 2010). Nevertheless, the gains in HRQoL from cCBT compared with control were small (Gerhards *et al.* 2010; Warmerdam *et al.* 2010) and not statistically significant, which is consistent with the analyses presented here (Gerhards *et al.* 2010; Hollinghurst *et al.* 2010; Warmerdam *et al.* 2010). Importantly, previous cost-effectiveness analyses have used cCBT effectiveness data from a developer-led trial where cCBT had clinical support by a practice nurse in contrast with the technical telephone support provided in REEACT (McCrone *et al.* 2004; Proudfoot *et al.* 2004; Kaltenthaler *et al.* 2006; National Institute for Health and Clinical Excellence, 2009). This may not be reflective of the type of support that would be feasible within the NHS and could have a considerable impact on the cost-effectiveness of cCBT, as clinical support has been shown to be a determinant of effectiveness for cCBT (Andersson & Cuijpers, 2009). Low adherence and engagement with cCBT in REEACT (less than 20% of patients on cCBT completed the treatment) (Gilbody *et al.* 2015; Littlewood *et al.* 2015) may explain the reduced effectiveness of the treatment when compared with the results of the developer-led trial where only 22% of patients on the cCBT arm withdrew from treatment (McCrone *et al.* 2004; Proudfoot *et al.* 2004).

It is important that any conclusions from these findings are assessed in relation to possible limitations. First, we have previously reported several possible limitations of the REEACT study, including: the selection of participants based on a definition of depression derived from a depression severity score as opposed to a structured diagnostic interview; insufficient statistical power to detect smaller effect sizes (not clinically significant) reported in entirely unsupported cCBT, despite large sample size, and potential crossover and dilution of effect (Gilbody *et al.* 2015). It is worth noting that statistical power to detect clinically significant improvements in depression does not necessarily translate into sufficient statistical power to detect differences in terms of cost-effectiveness, given the high variability of costs (Gray *et al.* 1997). In addition, it is possible that the follow-up period was insufficient to demonstrate the long-term benefits of cCBT. For the purposes of cost-effectiveness analyses, it is important to consider the time-frame over which costs and benefits are likely to differ between the interventions under consideration and in some instances these differences may need to be accounted for over a patient's lifetime. However, given the lack of difference in costs and QALYs between the arms during the trial, there appears to be no basis for inferring that any differences might occur in the future and therefore that conclusions might be altered if extrapolation was conducted. We also note that with reference to NICE guidance, the participants in the REEACT mostly had mild to moderately severe depression, but that some also had more severe disorder. NICE specifically recommends cCBT for lower-severity disorders, but in this pragmatic trial it was offered by GPs to people with a greater range of depression severity.

In conclusion, our findings suggest that technically supported cCBT programmes do not appear any more cost-effective than usual GP care alone for the management of depression in a primary care setting. Our results also suggest that a commercially developed programme appears no more cost-effective than a free-to-use cCBT programme.

Supplementary material

The supplementary material for this article can be found at https://doi.org/10.1017/S0033291717000289

Acknowledgements

We would like to extend our thanks to Thomas Devlin (University of York), Angela Swallow (University of Manchester) and Lone Gale (University of Bristol) for their valuable work in the collection of GP data. This work was funded by the UK National Institute for Health Research Health Technology Assessment Programme (project number 06/43/05).

The views and opinions expressed are those of the authors and do not necessarily reflect those of the Health Technology Assessment Programme, National Institute for Health Research, NHS or the Department of Health.

Declaration of Interest

None.

References

Andersson G, Cuijpers P (2009). Internet-based and other computerized psychological treatments for adult

depression: a meta-analysis. *Cognitive Behaviour Therapy* **38**, 196–205.

Barber J, Thompson S (2004). Multiple regression of cost data: use of generalised linear models. *Journal of Health Services Research and Policy* **9**, 197–204.

Bower P, Gilbody S (2005). Stepped care in psychological therapies: access, effectiveness and efficiency. Narrative literature review. *British Journal of Psychiatry* **186**, 11–17.

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

Brazier J, Roberts J, Tsuchiya A, Busschbach J (2004). A comparison of the EQ-5D and SF-6D across seven patient groups. *Health Economics* **13**, 873–884.

Brazier J, Usherwood T, Harper R, Thomas K (1998). Deriving a preference-based single index from the UK SF-36 health survey. *Journal of Clinical Epidemiology* **51**, 1115–1128.

Briggs A, Claxton K, Sculpher M (2006). *Decision Modelling for Health Economic Evaluation*. Oxford University Press: Oxford.

Brooks R (1996). EuroQol: the current state of play. *Health Policy* **37**, 53–72.

Christensen H, Griffiths KM, Jorm AF (2004). Delivering interventions for depression by using the Internet: randomised controlled trial. *BMJ* **328**, 265.

Curtis L (2012). *Unit Costs of Health and Social Care 2012*. Personal Social Services Research Unit (PSSRU), the University of Kent: Canterbury.

Department of Health (2012). National Schedules of Reference Costs: NHS Own Costs. NHS Reference Costs: Financial Year 2011 to 2012. Department of Health (https://www.gov.uk/government/publications/nhs-reference-costs-financial-year-2011-to-2012). Accessed February 2017.

Dolan P, Gudex C, Kind P, Williams A (1995). A social tariff for EuroQoL: results from a UK general population survey. In *CHE Discussion Paper*. Centre for Health Economics: York.

Fenwick E, Claxton K, Sculpher M (2001). Representing uncertainty: the role of cost-effectiveness acceptability curves. *Health Economics* **10**, 779–787.

Gerhards S, de Graaf LE, Jacobs L, Severens J, Huibers M, Arntz A, Riper H, Widdershoven G, Metsemakers J, Evers S (2010). Economic evaluation of online computerised cognitive–behavioural therapy without support for depression in primary care: randomised trial. *British Journal of Psychiatry* **196**, 310–318.

Gilbody S, Littlewood E, Hewitt C, Brierley G, Tharmanathan P, Araya R, Barkham M, Bower P, Cooper C, Gask L, Kessler D, Lester H, Lovell K, Parry G, Richards DA, Andersen P, Brabyn S, Knowles S, Shepherd C, Tallon D, White D; REEACT Team (2015). Computerised cognitive behaviour therapy (cCBT) as treatment for depression in primary care (REEACT trial): large scale pragmatic randomised controlled trial. *BMJ* **351**, h5627.

Gray AM, Marshall M, Lockwood A, Morris J (1997). Problems in conducting economic evaluations alongside clinical trials. Lessons from a study of case management for people with mental disorders. *British Journal of Psychiatry* **170**, 47–52.

Hollinghurst S, Peters TJ, Kaur S, Wiles N, Lewis G, Kessler D (2010). Cost-effectiveness of therapist-delivered online cognitive–behavioural therapy for depression: randomised controlled trial. *British Journal of Psychiatry* **197**, 297–304.

Joint Formulary Committee (2013). *British National Formulary*. BMJ Group and Pharmaceutical Press: London.

Kaltenthaler E, Brazier J, De Nigris E, Tumur I, Ferriter M, Beverley C, Parry G, Rooney G, Sutcliffe PA (2006). Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. *Health Technology Assessment* **10**, iii, xi–xiv, 1–168.

Kroenke K, Spitzer RL, Williams JBW (2001). The PHQ-9. *Journal of General Internal Medicine* **16**, 606–613.

Kupfer DJ, Frank E, Wamhoff J (1996). Mood disorders: update on prevention of recurrence. In *Interpersonal Factors in the Origin and Course of Affective Disorders* (ed. C Mundt and MJ Goldstein), pp. 289–302. Gaskell/Royal College of Psychiatrists: London.

Littlewood E, Duarte A, Hewitt C, Knowles S, Palmer S, Walker S, Andersen P, Araya R, Barkham M, Bower P, Brabyn S, Brierley G, Cooper C, Gask L, Kessler D, Lester H, Lovell K, Muhammad U, Parry G, Richards DA, Richardson R, Tallon D, Tharmanathan P, White D, Gilbody S; REEACT Team (2015). A randomised controlled trial of computerised cognitive behaviour therapy for the treatment of depression in primary care: the Randomised Evaluation of the Effectiveness and Acceptability of Computerised Therapy (REEACT) trial. *Health Technology Assessment* **19**, viii, xxi–171.

Manca A, Hawkins N, Sculpher MJ (2005). Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. *Health Economics* **14**, 487–496.

Matthews JN, Altman DG, Campbell MJ, Royston P (1990). Analysis of serial measurements in medical research. *BMJ* **300**, 230–235.

McCrone P, Knapp M, Proudfoot J, Ryden C, Cavanagh K, Shapiro DA, Ilson S, Gray JA, Goldberg D, Mann A (2004). Cost-effectiveness of computerised cognitive–behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *British Journal of Psychiatry* **185**, 55–62.

Moussavi S, Chatterji S, Verdes E, Tandon A, Patel V, Ustun B (2007). Depression, chronic diseases, and decrements in health: results from the World Health Surveys. *Lancet* **370**, 851–858.

National Institute for Health and Care Excellence (2013). *Guide to the Methods of Technology Appraisal*. NICE: London.

National Institute for Health and Clinical Excellence (2009). *Depression: The Treatment and Management of Depression in Adults* (*Update*). *Clinical Guideline 90*. National Institute for Health and Clinical Excellence: London.

Proudfoot J, Ryden C, Everitt B, Shapiro DA, Goldberg D, Mann A, Tylee A, Marks I, Gray JA (2004). Clinical efficacy of computerised cognitive–behavioural therapy for anxiety and depression in primary care: randomised controlled trial. *British Journal of Psychiatry* **185**, 46–54.

Roth A, Fonagy P (2005). *What Works for Whom: A Critical Review of Psychotherapy Research*. The Guilford Press: New York.

Royston P (2004). Multiple imputation of missing values. *Stata Journal* **4**, 227–241.

Spek V, Nyklícek I, Smits N, Cuijpers P, Riper H, Keyzer J, Pop V (2007). Internet-based cognitive behavioural therapy for subthreshold depression in people over 50 years old: a randomized controlled clinical trial. *Psychological Medicine* **37**, 1797–1806.

The EuroQol Group (1990). EuroQol – a new facility for the measurement of health-related quality of life. *Health Policy* **16**, 199–208.

Üstün TB, Ayuso-Mateos JL, Chatterji S, Mathers C, Murray CJL (2004). Global burden of depressive disorders in the year 2000. *British Journal of Psychiatry* **184**, 386–392.

Warmerdam L, Smit F, van Straten A, Riper H, Cuijpers P (2010). Cost-utility and cost-effectiveness of Internet-based treatment for adults with depressive symptoms: randomized trial. *Journal of Medical Internet Research* **12, e53**.

https://doi.org/10.1017/S0033291717000289 Published online by Cambridge University Press