Cost Effectiveness of Digital Interventions for Generalised Anxiety Disorder: A Model-Based Analysis

Dina Jankovic[1] · Pedro Saramago Goncalves[1] · Lina Gega[2] · David Marshall[3] · Kath Wright[3] · Meena Hafidh[1] · Rachel Churchill[3] · Laura Bojke[1]

Abstract

Background Digital interventions (DIs) are increasingly being used in mental health care, despite limited evidence regarding their value for money. This study aimed to evaluate the cost effectiveness of DIs for generalised anxiety disorder (GAD), in comparison with alternative care options, from the perspective of the UK health care system.

Methods An open-source decision analytic cohort model was used to extrapolate the results of a network meta-analysis over a patient's lifetime and estimate the costs and outcomes (quality-adjusted life-years) of DIs and their comparators. The net monetary benefit (NMB) and probability of cost effectiveness was estimated for each comparator, and we conducted a Value of Information analysis to evaluate the scale and drivers of uncertainty.

Results DIs were associated with lower NMB compared with medication and with group therapy, but greater NMB compared with non-therapeutic controls and with usual care. DIs that were supported by a clinician, an assistant or a lay person had higher delivery costs than purely patient-self-directed DIs, yielding a greater NMB when opportunity cost was above £3000/QALY. There was considerable uncertainty in the findings driven largely by uncertainty in the estimated treatment effects. The value of further research to establish the effectiveness of DIs for GAD was substantial, at least £12.9 billion.

Conclusions The high uncertainty about these results does not allow for recommendations based on the cost effectiveness of DIs. However, the analysis highlights areas for future research, and demonstrates that apparent cost savings associated with DIs can be offset by reduced effectiveness.

1 Background

Digital interventions (DIs) use software programmes accessed via computers, smartphones, audio-visual equipment and other devices to deliver therapeutic activities that aim to prevent and improve health problems. DIs lend themselves well to mental health care, where they have been used as alternatives or as add-ons to conventional therapies to improve access, patient choice and clinical outcomes. In England, National Health Service (NHS) investment in DIs is growing [1]. Such investment can be large and irreversible (e.g. through investment in training and infrastructure to deliver such interventions), so we need to understand the circumstances in which DIs offer value for money relative to alternative care options.

Generalised anxiety disorder (GAD) is a noteworthy condition for which DIs might be used to improve outcomes. It is the most common mental health problem in terms of weekly prevalence, but it is often misunderstood for other conditions (especially depression or panic) when self-reported, or it is mixed with other common mental disorders in research trials and evidence syntheses [2]. GAD can be magnified at times of crisis as it is defined by stress and worry about day-to-day life and struggling to tolerate uncertainty [3]; it is also associated with significant physical symptoms that lead to high health care and medical costs [4].

Compared with depression, evidence about the costs and outcomes of DIs specific to GAD populations is limited. Previous studies that evaluated the cost effectiveness of DIs with GAD populations [5, 6] compared a specific DI with usual care or individual therapy. To our knowledge, there are no studies that have synthesised the available evidence on all DIs for GAD in order to evaluate their cost effectiveness across different technologies and therapeutic modalities. Reviews of cost effectiveness of DIs were for mixed mood and anxiety disorders, so they were dominated by DIs for depression and did not report outcomes for GAD populations separately [7, 8].

✉ Dina Jankovic
  dina.jankovic@york.ac.uk

Extended author information available on the last page of the article

Key Points for Decision Makers

In mental health care, there is growing interest in using digital interventions as alternatives or as add-ons to conventional therapies to improve access, patient choice and clinical outcomes. However, digital interventions (DIs) may not be cost effective, as the apparent cost savings associated with cheaper delivery can be offset by reduced effectiveness.

In order to ensure DIs provide good value for money, further research is required to establish the effectiveness and the optimum role of digital interventions in the treatment of generalised anxiety disorder.

This study aims to evaluate the cost effectiveness of DIs, across different technologies and therapeutic modalities, in comparison with (a) conventional therapy (without any digital components), (b) medication, (c) non-therapeutic controls and (d) usual care, from the perspective of the UK's health care system.

2 Methods

2.1 Classification of Digital Interventions (DIs) and Their Alternatives

To pool and compare different types of DIs and their controls for GAD, a systematic literature review and network meta-analysis [9] was conducted in which DIs and their alternatives were classified according to three criteria: (a) whether they were a psychological/behavioural intervention or a non-therapeutic psychological/behavioural control; (b) whether they were digital or non-digital; and (c) whether they were supported.

'Digital' infers a software-based platform bespoke to the delivery of a specific activity. 'Controls' could be a psychological placebo (e.g. 'sham' activity), an attention control (e.g. non-therapeutic interaction with a researcher), or a change in usual care due to research processes (e.g. regular monitoring to ensure retention to follow-up). 'Supported' interventions/controls were defined by having a two-way interaction between the patient and another person (clinician or lay person).

Waiting lists and usual care were classified under usual care unless an active component (e.g. monitoring, sham activity) was introduced, in which case the waiting list/usual care was classified as non-therapeutic control. An additional classification group was included for pharmacological interventions (i.e. medication).

In total, we included seven comparators based on our classification criteria: medication, where selective serotonin reuptake inhibitors (SSRIs) were the only pharmacotherapy identified in the review that has been directly compared with digital interventions for GAD; face-to-face group therapy, the only supported non-digital intervention identified in the review; supported digital intervention (SDI); unsupported digital intervention (UDI); supported digital control (SDC); unsupported digital control (UDC); usual care. There were no available clinical studies that compared digital interventions with unsupported non-digital interventions or unsupported non-digital controls, and no studies with supported non-digital controls that used GAD-7 as an outcome. Details of the digital interventions and controls are provided in Supplementary File 1 of the Electronic Supplementary Material (ESM). It is possible that the taxonomy did not account for all differences between interventions within each group. Alternative, more granular classification was explored in the network meta-analysis [9], and is discussed in further detail in Sect. 4, under limitations.

2.2 Analytical Perspective

Analysis was conducted from the NHS and Personal Social Services perspective. A full incremental analysis was undertaken comparing all seven interventions/controls simultaneously over a patient's lifetime. The cost-effectiveness analysis methods followed the National Institute for Health and Care Excellence (NICE) reference case [10]. We included costs associated with health care use and intervention delivery, as incurred by the health system. Outcomes were measured in terms of quality-adjusted life-years (QALYs) that capture the length and the health-related quality of life (HRQoL).

2.3 Model Structure

The model structure (Fig. 1) was based on previously published models, namely those by Kumar et al. [5], adapted to be used with the available data on outcomes in this study. The model was validated by clinical researchers on the team (LG) and advisors to the project. In the model, patients' outcomes were derived from GAD severity determined by the Generalised Anxiety Disorder-7 item questionnaire (GAD-7), whose scores denoted no anxiety (scores 0–4), mild (5–9), moderate (10–14) or severe (15–21) anxiety [11]. Each health state (GAD severity level) was associated with a corresponding health care cost, HRQoL and mortality, driven by GAD and associated comorbidities.

At the start of the model, patients were in one of three health states: mild, moderate or severe anxiety. At each subsequent cycle of the model (every 3 months over patients' lifetime), they could remain in this health state or transition to a better or worse health state, including no GAD. The distribution of patients across the health states was derived
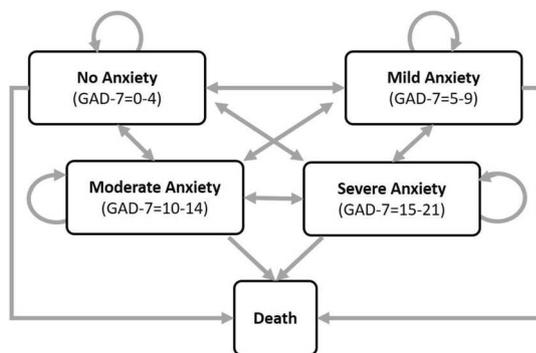
Fig. 1 Model structure

from the distribution of the expected GAD-7 score at each time point. As patients passed through states they accrued costs and QALYs. The effect of DIs is to reduce the severity of GAD, therefore reduce the probability of being in moderate and severe GAD states, with an associated impact on costs and QALYs. Patient were 46.7 years old and female on entering the model; the patient population was based on the population used to inform state-specific utilities in the model where 72.4% of the population was female [12]. The model was implemented in RStudio, version 1.3.1093 [13], the code is available in the online supplementary files (see ESM).

2.4 Model Parameters

2.4.1 Intervention Effectiveness

Changes in GAD-7 scores, as a result of the seven comparators, were informed by a network meta-analysis of DIs for patients diagnosed with GAD [9]. The meta-analysis reported GAD-7 scores after treatment (3–12 weeks), adjusted by baseline scores. GAD-7 scores were modelled as a continuous variable; scores generated in the meta-analysis model that fell between severity states (e.g. a GAD-7 score of 9.5), were rounded up.

Baseline GAD-7 scores were used to determine the proportion of patients in the three health states (mild, moderate, severe GAD). Post-treatment (cycle 2) patients can transition to another state, including no GAD and death, or remain in their existing state. Changes in GAD severity in the remaining cycles were estimated based on evidence from the literature. Without treatment, patients' GAD symptoms were assumed to improve over time, as reported by Yonkers et al. [14]. Specifically, 15% of patients recovered in the first year, a further 10% in the second year, and a further 5% in the third year. In the model, recovery was defined as a move from mild, moderate or severe anxiety to the next (lower) anxiety state. In the base case, the treatment effect

was assumed to remain constant indefinitely, implying that, while patients' GAD symptoms may change over time, the proportion of patients in each health state, on average, would not change. The cost of any additional treatment was assumed to be captured in health care costs. Alternative, more conservative assumptions were explored in sensitivity analysis (Sect. 2.5.2). In addition, in the base case, patients on treatment were assumed to improve over time at the same rate as those who hadn't received treatment.

2.4.2 State-Specific Utilities and Costs

Targeted searches were conducted to inform state-specific health care costs and utilities in no, mild, moderate and severe anxiety. Details of the search strategies and results are provided in Supplementary File 2 (see ESM).

In the base case, the utilities were informed by Revicki et al. [12], a study conducted in 297 adult patients with GAD recruited from an integrated care setting in the US (72%, mean age 47.69 years). In the study, GAD severity states were defined by the Hamilton Anxiety Scale (HAM-A) [15] score, an assessment tool highly correlated with the GAD-7 questionnaire ($r = 0.852$ [16]). Utilities were elicited using an SF-12 Health Survey (SF-6D).

The HRQoL/utility scores associated with the GAD-7 states were assumed to follow an underlying age depreciation over time, according to UK population utility score norms (per year of age) [17].

State-specific health care costs were informed by a UK-based study [18], adjusted to 2019 GBP using the overall Consumer Price Index [19, 20].

Costs and benefits were discounted at 3.5% over the time horizon of the model [10, 21].

2.4.3 Mortality

Age- and sex-dependent mortality risk in patients with no anxiety was obtained from the Office for National Statistics [22]. Excess mortality in mild, moderate and severe anxiety states was derived from Michal et al. [23], who reported all-cause excess mortality in patients with mild, moderate and severe anxiety or depression, defined by PHQ-4 scores. All excess mortality was assumed to be caused by GAD, capturing deaths due to suicide and GAD-related co-morbidities. The reported risk ratios were applied to all-cause mortality.

2.4.4 Intervention Cost

Intervention costs were derived from published literature. The cost of usual care was assumed to depend on GAD severity and was captured in state-specific health care costs (Sect. 2.4.2). The cost of pharmacotherapy was assumed to represent the cost of medication (£16.42 representing

the mean cost of all SSRIs [24] weighted by the volume dispensed in January 2020 [25]), the dispensing fees (12 prescriptions dispensed annually at £1.26 each [24]), and GP appointments (7.5 in the first year of treatment and 4 thereafter for a further 4 years, as per NICE guidance for prescribing [26], at £42.60 each). On advice from clinical researchers on the project, medication was assumed to be prescribed for 5 years as a conservative estimate to prevent underestimating its cost. The duration was chosen because antidepressant medication in treatment-naïve patients may be given for up to 2 years, but relapse is common within the first year after of stopping the medication, which may lead to another 2 years of prescription.

The cost of face-to-face group therapy was based on the time spent with a therapist, multiplied by the cost of the therapist's time (£53 per hour [27]). The group sizes were 5–9, lasting 9–10.5 h [28, 29]. The intervention was thus costed at 1.5 h of therapist time per patient (10.5 h/7 patients).

A recent review of economic evaluations of digital mental health interventions [30] found that the methods for costing these interventions varied greatly between studies. In this study, the cost of the digital component of interventions/controls and their maintenance was assumed to be zero in the base case, assuming that, if the intervention is rolled out nationally, the marginal cost per patient would be negligible. Alternative costs were explored in scenario analyses (Sect. 2.5.2). Unsupported digital interventions and controls (UDCs and UDIs) were assumed to incur no additional human resource costs for their delivery. The human resource cost of SDIs and SDCs was calculated separately, based on the level of support (e.g. who provided the support, for how long, and how often) typically required to deliver them, detailed in supplementary file 3 (see ESM).

2.5 Cost-Effectiveness Analysis

Cumulative costs and QALYs over a lifetime were used to derive the net monetary benefit (NMB) for each comparator, using Eq. (1). Increasing NMB implies better value for money.

$$NMB = QALY \times k - cost \tag{1}$$

where k is the opportunity cost of the health system [31].

NMB is the preferred method for comparing cost effectiveness of multiple treatment options [32]; it represents the difference between the benefit incurred by an intervention and its opportunity cost [33]. In our base case, k was £15,000/QALY, comparable to the empirical estimate of the opportunity cost in the UK [34], but alternative values (£0/QALY to £30,000/QALY) were explored in sensitivity analysis.

2.6 Sensitivity Analysis

2.6.1 Probabilistic Analysis

Probabilistic analysis was conducted to characterise the uncertainty associated with model parameters, and determine their impact on cost effectiveness. Each parameter was sampled 25,000 times from its probability distribution to match the number of iterations output from the NMA that directly informed intervention effectiveness in the economic model. The model parameters and their probability distributions are shown in Table 1 and Supplementary File 4 (see ESM).

2.6.2 Deterministic Scenario Analysis

One-way scenario analysis was performed to evaluate the sensitivity of the model results to our assumptions. The following scenarios were explored regarding the GAD trajectory, health care resource use and intervention costs:

Five additional scenarios regarding the GAD trajectory, assuming different rates of diminishing treatment effect and assuming no spontaneous improvement. Specifically, the treatment effect was assumed to remain constant for 1 year, then either diminish immediately or diminish gradually for 10 years before returning to pre-treatment GAD-7 scores. The scenarios were chosen based on clinical opinion (see Supplementary File 5 in the ESM for further detail).
Two alternative scenarios regarding health care resource use, where state-specific health care costs were informed using alternative studies: Vera-Llonch et al. [35] and Kumar et al. [5] (see Supplementary File 6 in the ESM for details).
Alternative costs of digital interventions, where a threshold analysis was performed to identify the maximum cost of DIs that would make them good value for money.
Alternative level of support in supported digital controls—5 minutes per patient, delivered by non-clinical staff (£1.50) or by clinical psychologists (£3.00).

2.7 Value of Information Analysis

The results from probabilistic sensitivity analysis were used to conduct value of information (VOI) analysis, on a population level. Estimated population-level value of perfect information (EVPI$_p$) represents the difference between the NMB derived from perfect information (where the optimum treatment is known with certainty) and existing information (i.e. the expected net benefit from the treatment that is the most cost effective under the current state of knowledge)

Table 1 Model parameters, with measures of uncertainty

Parameter	Mean	Probability distribution	Source
Baseline GAD-7 score	10.68	~ N (10.68, 1.83)	Saramago Goncalves et al. submitted [9]
GAD-7 score after treatment			Saramago Goncalves et al. submitted [9]
UDC	8.77	~ N (8.77, 3.72)	
SDC	8.76	~ N (8.76, 3.14)	
UDI	8.01	~ N (8.01, 3.16)	
SDI	7.22	~ N (7.22, 1.82)	
Group therapy	6.51	~ N (6.51, 3.59)	
SSRIs	4.08	~ N (4.08, 1.89)	
Interventions cost			
Usual care, UDC, UDI	£0		Assumed
SDC	£18		Derived
SDI	£80		Derived
Group therapy	£80		Derived
SSRIs (per year)	£351.04 in year 1, £201.94 thereafter		Derived
Health care cost		Gamma (shape, scale)	Kaltenthaler et al. [18]
No anxiety	86	(1.960, 0.016)	
Mild anxiety	200	(0.848, 0.003)	
Moderate anxiety	210	(0.295, 0.001)	
Severe anxiety	324	(0.320, 0.0007)	
Dead	0		
Utilities		Mean (SD) used to derive beta parameters*	Revicki et al. [12]
No anxiety	0.72	0.72 (0.10)	
Mild anxiety	0.64	0.64 (0.10)	
Moderate anxiety	0.60	0.60 (0.10)	
Severe anxiety	0.53	0.60 (0.10)	
Dead	0		
Age-related utility decrements	Age specific, see Table S3.3 in ESM		
Mortality—general population	Sex, and age specific, see Table S3.3 in ESM		
Excess mortality (RR)		Non-parametric CI**	Michal et al. [23]
Mild anxiety	1.20	(0.82–1.67)	
Moderate anxiety	1.58	(1.06–2.24)	
Severe anxiety	2.17	(1.47–3.05)	

CI confidence interval, *ESM* electronic supplementary material, *GAD-7* generalised anxiety disorder 7 score, *RR* relative risk, *SDC* supervised digital controls, *SDI* supervised digital interventions, *SSRIs* selective serotonin reuptake inhibitors, *UDC* unsupervised digital controls, *UDI* unsupervised digital interventions
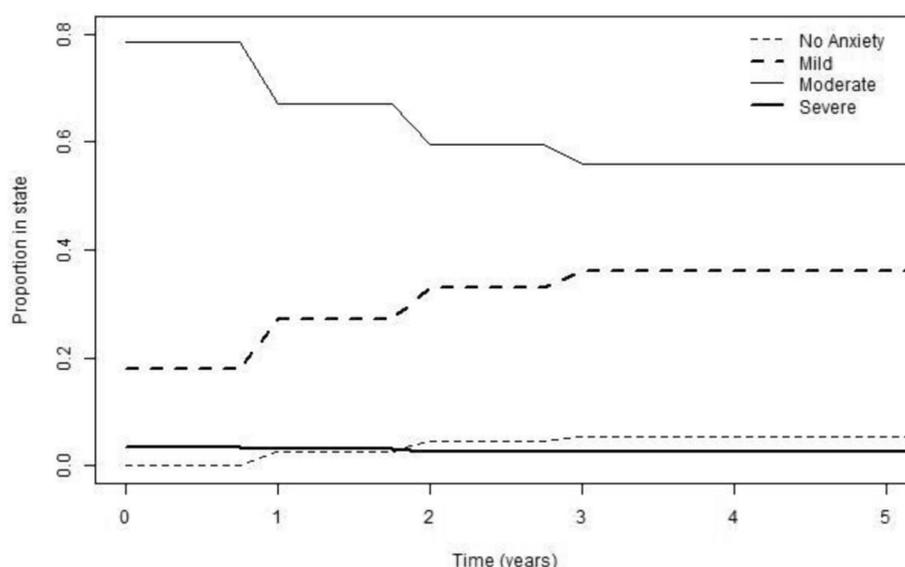
*Uncertainty in utilities was derived by fitting a beta distribution to the reported mean and standard error, using the method of moments [4]

**Beta distribution applied to the risk of death in each state, then 25,000 random samples used to derive the relative risk

[36]. Intuitively, it is interpreted as the maximum amount that should be spent on research to resolve model uncertainty, given the potential impact, or benefit, of the research findings. To derive EVPI$_P$, the incidence of GAD in England was assumed to be 250,000 people per annum (assuming incidence is 4.9% per annum [37], and 10% of patients receive the intervention), and the lifetime of the intervention was conservatively estimated to be 5 years. Methods are described in further detail in Supplementary File 7 (see ESM).

The estimated value of perfect parameter information at population level (EVPPI$_P$) was derived to identify the parameters in the model that drove uncertainty relevant to the adoption decision. EVPPI$_P$ represents the maximum amount that should be spent on research to resolve uncertainty around an individual parameter (or group of parameters), given the potential impact, or benefit, of the research findings. EVPPI$_P$ was derived using the non-parametric method developed by Strong et al. [38], using the SAVI interface (http://savi.shef.ac.uk/SAVI/). EVPPI$_P$ was

Fig. 2 Movement through health states for the initial 5 years without treatment

3 Results

3.1 Generalised Anxiety Disorder Trajectory

Figure 2 shows the proportion of patients in each health state for the initial 5 years, with usual care. At the start of the model, the majority of patients had mild or moderate GAD (17.9% and 78.6%, respectively). In the base case, patients' GAD was assumed to improve spontaneously for the first 3 years of the model [13], resulting in a decrease in the proportion who have moderate and severe GAD, and an increase in no or mild GAD. After 3 years the proportions remained constant in the living population. The proportion in each anxiety state after treatment is shown in Supplementary File 4, Table S4.1 (see ESM).

The reduction in GAD-7 scores after receiving treatment, for the seven comparators, is shown in Fig. 3. The initial reduction in GAD-7 reflects outcomes, as determined by the meta-analysis by Saramago Goncalves et al. [9]. The effect of each comparator was assumed to remain constant indefinitely and, after 1 year, GAD-7 scores further decreased as patients' symptoms continued to improve at the same rate as without treatment. However, the GAD-7 reduction decreased as, following treatment, there were fewer patients who could recover spontaneously, compared with no treatment.

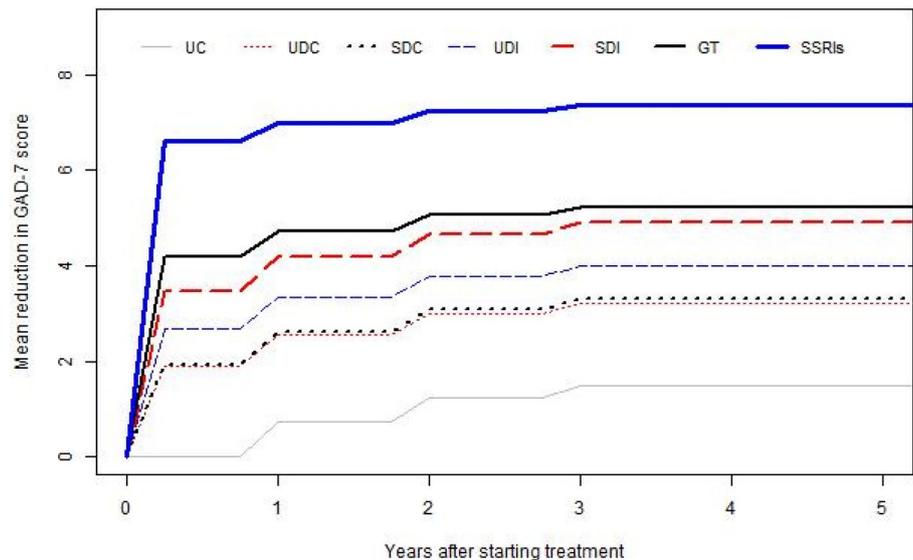### 3.2 Costs and Outcomes of DIs Compared with Alternatives

Table 2 shows the costs and outcomes associated with each of the seven comparators. The QALY and LY gains reflect the results from the meta-analysis [9]: on average, SSRIs were associated with lower anxiety scores post-treatment, followed by face-to-face group therapy, then by SDIs, and then UDIs. Both SDIs and UDIs were associated with lower GAD-7 scores post-treatment compared with SDCs and UDCs, and with no treatment. Differences between comparators were small and uncertain, reflected in wide and overlapping confidence intervals.

Differences in health care costs largely followed the reverse order of QALY gains—health care costs were the lowest for patients taking medication, and the highest in patients who received usual care. This is because increasing GAD severity is associated with an increase in health care costs (see Table 1). Health care costs were highly uncertain with overlapping confidence intervals.

SDIs and face-to-face group therapy were associated with the same intervention costs, as they include the same level of human resources. Treatment with SSRIs was more expensive, as it requires contact with health care professionals for 5 years, for the duration of treatment. The total costs of different comparators follow the same order as health care costs, except SDIs, which incur higher total costs than UDIs due to the higher intervention cost.

Table 2 shows the incremental costs and effects, and the NMB of all seven comparators, while Fig. 4 shows the cost-effectiveness acceptability curves (CEACs) for each comparator for different opportunity costs. Net benefit and the probability of cost effectiveness in Table 2 are shown for two opportunity costs: £0/QALY representing a decision

The reduction in GAD-7 scores after receiving treatment, for the seven comparators, is shown in Fig. 3. The initial reduction in GAD-7 reflects outcomes, as determined by the

derived for five groups of parameters: treatment effect (post-treatment GAD-7 scores, seven parameters); state-specific health care costs (four parameters); state-specific utilities (four parameters); excess death (three parameters—relative risk of death in mild, moderate and severe anxiety), age-related utility decrements (54 parameters, utility decrements every year until all patients in the model die). For methods, see Supplementary File 7 in the ESM.

Fig. 3 Change in mean GAD-7 score for the initial 5-year period after starting treatment

UC=Usual care; UDC=Unsupported Digital Control; SDC=Supported Digital Control; UDI=Unsupported Digital Intervention; SDI=Supported Digital Intervention; GT = group therapy; SSRIs=Selective Serotonin Reuptake Inhibitors.

maker who will only implement cost-saving interventions, and £15,000/QALY close to the empirical estimate of the opportunity cost in England [34]. The results at opportunity costs > £15,000/QALY are not shown as they did not change significantly (see Fig. 4).

SSRIs resulted in a dominant ICER and highest NMB at all opportunity costs, followed by group therapy, as both led to lower total costs and greater QALY gains than DIs. Usual care had the lowest NMB. Digital interventions (UDIs and SDIs) had a higher NMB than digital controls (UDC and SDC). SDIs had greater QALY gains but also higher costs than UDIs, so their NMB depended on the opportunity cost of the health system. When the opportunity cost is £0/QALY, UDIs are ranked above SDIs, while the opposite is the case when opportunity cost increases to £15,000/QALY.

Results are uncertain, reflected in the wide confidence intervals associated with NMB, and the probability of each intervention being the most cost effective in Table 2. For example, while SSRIs have the highest NMB, there is a high probability it is not the most cost-effective comparator (0.949 probability of not being cost effective at £0/QALY opportunity cost, and 0.431 at £15,000/QALY).

Accumulated costs and outcomes over time are shown in Supplementary File 8 (see ESM). The longer the analysis time horizon the greater the differences in the QALY gain, as the benefits accrue over time. Differences in health care cost are driven by the clinical effectiveness of interventions, and so, as for QALY gains, the longer the time horizon the greater the cost differences. Treatment with SSRIs is the most expensive in the short term, but as health care cost savings accrue, and treatment cost reduces after 5 years, the total cost for SSRIs increases at a lower rate than for other comparators, eventually becoming the cheapest treatment option, after face-to-face group therapy.

3.3 Value of Information (VoI) Analysis

In the base case, the EVPI$_p$ was £16.2 billion, indicating high uncertainty in the model. EVPI$_p$ at a range of opportunity costs is shown in supplementary file 9 (see ESM), with the lowest value of £11.4 billion when opportunity cost is £4000/QALY. EVPPI$_p$ analysis suggested that uncertainty in the treatment effect had the greatest impact on the model results; the estimated VoI about the treatment effect was £12.9 billion.

3.4 Scenario Analyses

Scenario analyses were performed to explore the sensitivity of the findings to assumptions made regarding the GAD trajectory, health care costs and interventions costs. The results were not sensitive to any of the alternative scenarios—the order of cost effectiveness did not change, only the magnitude of the difference. SSRIs dominate all other comparators, while SDIs are more effective and costlier than UDIs (see supplementary file 10 in the ESM for details). Since DIs were associated with higher health care costs than group therapy and SSRIs, exploring the effect of increasing their cost even further was unnecessary.

Table 2 Mean total cost and effect of digital interventions (DIs) and alternatives for generalised anxiety disorder (GAD) (95% confidence interval)

	SSRIs	Group therapy	SDI	UDI	SDC	UDC	UC
QALYs	12.9 (11.5–14.3)	12.2 (10.4–14.2)	11.9 (10.5–13.7)	11.8 (9.7–14.1)	11.6 (9.2–14.0)	11.5 (8.9–14.1)	11.1 (9.4–12.4)
Life-years	37.2 (34.1–38.6)	36.4 (32.09–38.6)	36.4 (32.5–39.0)	36.0 (31.4–38.8)	35.7 (31.0–38.7)	35.6 (30.5–38.6)	34.9 (30.9–38.6)
Health care cost (£)	10,640 (725–45,948)	12,362 (49–59,092)	14,218 (139–63,152)	14,114 (25–68,008)	14,605 (13–74,297)	14,822 (18–75,923)	16,069 (1–87,123)
Intervention cost (£)	1115 (1113–1115)	80	80	0	18	0	0
Total cost (£)	11,754 (1839–47,063)	12,442 (128–59,171)	14,298 (218–63,232)	14,114 (25–68,008)	14,623 (31–74,315)	14,822 (18–75,923)	16,069 (1–87,123)
Incremental cost compared with SSRIs (£)		688 (− 38,229 to 49,867)	2543 (− 36,862 to 51,940)	2360 (− 37,391 to 57,798)	2868 (− 37,785 to 63,454)	3067 (− 38,072 to 65,451)	4315 (− 38,184 to 75,859)
Incremental QALYs compared with SSRIs		− 0.7 (− 3.4 to 2.2)	− 1.0 (− 3.0 to 0.8)	− 1.2 (− 3.6 to 1.8)	− 1.4 (− 3.8 to 1.4)	− 1.4 (− 4.5 to 1.8)	− 1.9 (− 3.9 to 0.3)
Mean ICER (£/QALY)	Dominant	Dominated	Dominated	Dominated	Dominated	Dominated	Dominated
k = £0/QALY							
NMB (£)	− 11,754 (− 47,063 to − 1839)	− 14,298 (− 63,232 to − 218)	− 14,298 (− 63,232 to − 218)	− 14,114 (− 68,008 to − 25)	− 14,623 (− 74,315 to − 31)	− 14,822 (− 75,923 to − 18)	− 16,069 (− 87,123 to − 1)
PCE	0.051	0.107	0.107	0.151	0.159	0.165	0.238
k = £15,000/QALY							
NMB (£)	181,975 (131,817–209,180)	164,551 (111,833–197,813)	164,551 (111,833–197,813)	162,334 (97,837–205,226)	158,586 (88,232–203,339)	158,335 (81,276–205,462)	149,671 (74,016–180,051)
PCE	0.431	0.070	0.070	0.092	0.066	0.091	0.015

ICER incremental cost-effectiveness ratio, *k* opportunity cost, *NMB* net monetary benefit, *PCE* probability of cost effectiveness, *QALY* quality-adjusted life-years, *SDC* supervised digital controls, *SDI* supervised digital interventions, *SSRIs* selective serotonin reuptake inhibitors, *UC* usual care, *UDC* unsupervised digital controls, *UDI* unsupervised digital interventions

4 Discussion

Interest in the use of DIs for the treatment of mental health disorders is growing, but it is not clear whether they represent value for money in the treatment of GAD, the most highly prevalent mental health problem. This study used a decision analytic model to evaluate the cost effectiveness of DIs, across different types of technology and therapeutic modalities, from the perspective of the UK health care system, compared with group therapy, SSRIs, non-therapeutic controls and usual care.

4.1 Key Findings

The expected net benefit was the highest for SSRIs followed by group therapy. All DIs and non-therapeutic controls led to higher net benefit than usual care. Supported DIs were associated with higher costs than unsupported DIs, but their NMB was greater when the opportunity cost was £3000/QALY or higher.

These results are highly uncertain, with the VoI estimated to be > £11 billion. The VoI represents not only the value of further research in order to resolve model uncertainty, but also the scale of the QALYs loss and incurred costs, if the decision about whether to fund DIs, based on existing evidence, is incorrect. The EVPPI$_P$ analysis found that uncertainty in the treatment effect had the greatest impact on the model results. The treatment effect was fundamental to establishing the cost effectiveness of DIs for GAD because the costs were driven largely by clinical outcomes (better GAD outcomes lead to lower total health care costs, compensating for higher intervention costs). The value of further research to establish the effectiveness of DIs for GAD is therefore substantial, approximated at £12.9 billion.

Fig. 4 Cost-effectiveness acceptability curves (CEACs) for each intervention and control
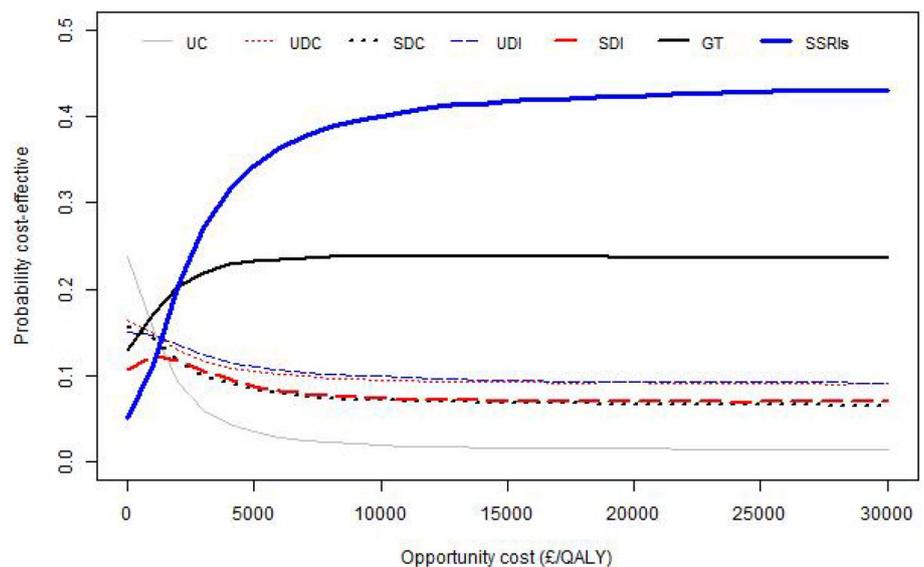
UC=Usual care; UDC=Unsupported Digital Control; SDC=Supported Digital Control; UDI=Unsupported Digital Intervention; SDI=Supported Digital Intervention; GT = group therapy; SSRIs=Selective Serotonin Reuptake Inhibitors.

4.2 Comparison with the Existing Literature

Previous studies evaluated the cost effectiveness of one specific DI for GAD using a single source for clinical data. Kumar et al. [5] evaluated the cost effectiveness of an SDI (mobile self-directed cognitive behaviour therapy—CBT) against individual CBT and usual care, and found the SDI to be cost saving against both comparators; however, the study used non-comparative data where the effect of mobile CBT was measured in a single-arm pilot study, while the effect of individual CBT was derived from a meta-analysis. Furthermore, the authors did not explore probabilistic uncertainty in the cost-effectiveness analysis. Dear et al. [6] evaluated the cost effectiveness of an SDI (internet-based self-directed CBT) compared with usual care, using outcomes from an RCT. The authors found that the SDI was more effective, but also costlier than usual care, with a high probability of being cost effective when the opportunity cost was AUS$40,000/QALY (~ £20,000/QALY at current exchange rate). Broadly, our analysis supports the findings by Kumar et al. [5] and by Dear et al. [6] that SDIs may be cost effective compared with usual care, albeit with great uncertainty. No clinical studies included head-to-head comparisons between DIs and individual CBT for GAD populations (only group CBT), so we could not comment on the finding by Kumar et al. [5] that DIs are cost effective compared with individual CBT.

4.3 Limitations

Estimates of clinical effectiveness used in the model were informed by pooled evidence for different but comparable interventions according to the taxonomy developed by Saramago Goncalves et al. [9]. It is possible that the taxonomy did not account for all differences between digital interventions, leading to further uncertainty in effectiveness estimates. However, Saramago Goncalves et al. [9] found that further granulation of intervention characteristics, to account for additional factors such as contact time, led to inconsistent effectiveness results with greater uncertainty, suggesting that there is no evidence that accounting for these additional characteristics could provide a better intervention taxonomy and reduce uncertainty in the findings. Furthermore, a similar approach to estimating the effectiveness of DIs has been taken in previous studies on depression [39, 40].

Limitations in data availability led to several assumptions regarding model parameters. Utilities and excess mortality in different severity states were informed by studies that used measures other than GAD-7, specifically HAM-A [12] and PHQ-4 [23]. HAM-A has been found to be highly correlated with GAD-7 [16], while correlation between PHQ-4 and GAD-7 is unclear. The cost of health care for different levels of GAD severity was based on research from 2005 (adjusted for inflation) [18] and on non-UK health care resource use in the scenario analyses [5, 35]. All three studies led to similar conclusions in terms of ranking and uncertainty in the cost

effectiveness of DIs and their alternatives. The trajectory of the treatment effect of different comparators is unclear. In our analysis we tested extreme scenarios, where the treatment effect lasts indefinitely (base case), and where it diminishes immediately after stopping treatment (scenarios 3 and 6 in supplementary file 5, see ESM) and found that the assumptions did not impact the conclusions of the cost-effectiveness analysis. When calculating costs, the duration of treatment with SSRIs was conservatively assumed to be 5 years. In practice, treatment is likely to last < 5 years, reducing the intervention cost. The lower intervention cost would not change the conclusion of the analysis, where total cost with SSRIs is already expected to be lower than with other comparators.

4.4 Policy Implications and Further Research

This study found that there is considerable uncertainty around the cost effectiveness of DIs for GAD, and that existing evidence is not sufficient to make an informed decision about whether these interventions represent good value for money.

However, the study helps identify key topics for further research, namely primary data collection (e.g. RCTs) to establish the clinical effectiveness of digital interventions in the GAD population. Firstly, the study found that clinical effectiveness was the key driver of cost effectiveness, as apparent cost savings associated with digital interventions (compared with face-to-face group therapy and SSRIs) can be offset by higher downstream health care costs due to reduced effectiveness, even if that reduction is small and statistically non-significant. Secondly, the uncertainty in the clinical effect was highlighted as the key driver of uncertainty in the cost-effectiveness model both by the VoI analysis in section 3.3 and by the results of the scenario analyses in supplementary file 10 (see ESM), where none of the scenarios had a significant impact on the model findings because the short-term effectiveness of the comparators was the key determinant of the cost effectiveness.

Primary data collection would also enable the analysis of how to best deliver digital interventions to GAD patients. Specifically, we identify four factors that could impact the cost effectiveness that were not explored in our analysis due to data sparsity.

a. Optimum design of digital interventions. Our study categorised digital interventions based on whether they included support and whether they included an active intervention. The categorisation was based on best available evidence at the time of the study. Further studies evaluating a variety of digital interventions in this population could inform how other factors, such as content and delivery mode, impact the effectiveness of digital interventions.

b. Heterogeneity in patient response. It is possible that digital interventions are more effective in some patient populations. Understanding patient characteristics that drive the effectiveness of digital interventions could allow more targeted treatment delivery.

c. Treatment sequencing. Our model compared the lifetime effect of single treatments for GAD, whereas in practice, patients can receive multiple cycles of the same therapy or a combination of therapies concurrently or in sequence. We did not model the cost effectiveness of sequential or combined treatments due to lack of data on effectiveness. As such, we do not know whether DIs may be cost effective as a first-line treatment in a stepped care model before medication and individual or group therapy is offered to those who do not respond to DIs.

d. Role of DIs in increasing capacity of the health system. The analysis does not consider the capacity of the health system to deliver treatment for GAD. Digital interventions may be used to improve capacity in health systems where there are significant waitlists for face-to-face treatment—a challenge likely to grow as result of increased incidence of mental health issues and decreased capacity for delivering face-to-face care associated with the COVID-19 pandemic. The health consequences of being able to offer treatment to a larger number of GAD patients, or reducing delay to treatment, may be substantial and may justify a lower effectiveness of DIs, as demonstrated in a recent study on depression [40].

5 Conclusions

This study is the first to collate all available clinical evidence to evaluate the cost effectiveness of DIs for GAD, irrespective of the type of DI used, and in comparison with multiple alternative care options. On average, DIs were associated with lower NMB compared with SSRIs and with face-to-face group therapy, but greater NMB compared with non-therapeutic controls and with usual care. Although supported DIs had a higher human resource cost than unsupported DIs, they led to greater NMB when opportunity cost was £3000/QALY or greater. The high uncertainty of the results prevents any firm conclusions about the cost effectiveness of DIs. However, the analysis highlights key areas for further research: primary data collection to establish the effectiveness and the optimum role of digital interventions in the treatment of GAD.

Declarations

References

1. NHS England, Five Year Forward View, 2014. https://england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf. Accessed 31 Jan 2018.

2. McManus S, Bebbington P, JenkinsR, and Brugha T. Adult psychiatric morbidity survey: survey of mental health and wellbeing, England, 2014. https://digital.nhs.uk/data-and-information/publications/statistical/adult-psychiatric-morbidity-survey/adult-psychiatric-morbidity-survey-survey-of-mental-health-and-wellbeing-england-2014. Accessed 19 May 2020.

3. World Health Organisation. ICD-11: International Classification of Diseases 11th Revision. The global standard for diagnostic health information, 2018.

4. Revicki DA, et al. Humanistic and economic burden of generalized anxiety disorder in North America and Europe. J Affect Disord. 2012;140(2):103–12.

5. Kumar S, Bell MJ, Juusola JL. Mobile and traditional cognitive behavioral therapy programs for generalized anxiety disorder: a cost-effectiveness analysis. PLoS ONE. 2018;13(1): e0190554.

6. Dear BF, et al. Clinical and cost-effectiveness of therapist-guided internet-delivered cognitive behavior therapy for older adults with symptoms of anxiety: a randomized controlled trial. Behav Ther. 2015;46(2):206–17.

7. Arnberg FK, Linton SJ, Hultcrantz M, Heintz E, Jonsson U. Internet-delivered psychological treatments for mood and anxiety disorders: a systematic review of their efficacy, safety, and cost-effectiveness. PLoS ONE. 2014;9(5): e98118.

8. Massoudi B, Holvast F, Bockting CLH, Burger H, Blanker MH. The effectiveness and cost-effectiveness of e-health interventions for depression and anxiety in primary care: a systematic review and meta-analysis. J Affect Disord. 2019;245:728–43.

9. Saramago Goncalves P, Gega L, Marshall D, Nikolaidis GF, Jankovic D, Melton H, Dawson S, Churchill R, Bojke L. Digital interventions for generalised anxiety disorder (GAD): systematic review and network meta-analysis. Front Psychiatry. 2021. https://doi.org/10.3389/fpsyt.2021.726222.

10. National Institute for Health and Care Excellence (NICE). Guide to methods of technology appraisal, 2013.

11. Spitzer RL, Kroenke K, Williams JBW, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. Arch Intern Med. 2006;166(10):1092–7.

12. Revicki DA, Brandenburg N, Matza L, Hornbrook MC, Feeny D. Health-related quality of life and utilities in primary-care patients with generalized anxiety disorder. Qual Life Res. 2008;17(10):1285–94.

13. RStudio Team (2020). RStudio: integrated development environment for R. RStudio, PBC, Boston, MA. http://www.rstudio.com/. Accessed June 2021.

14. Yonkers KA, Warshaw MG, Massion AO, Keller MB. Phenomenology and course of generalised anxiety disorder. Br J Psychiatry. 1996;168:308–13.

15. Maier W, Buller R, Philipp M, Heuser I. The Hamilton Anxiety Scale: reliability, validity and sensitivity to change in anxiety and depressive disorders. J Affect Disord. 1988;14(1):61–8.

16. Ruiz MA, Zamorano E, García-Campayo J, Pardo A, Freire O, Rejas J. Validity of the GAD-7 scale as an outcome measure of disability in patients with generalized anxiety disorders in primary care. J Affect Disord. 2011;128(3):277–86.

17. Kind P, Hardman G, Macran S. UK Population Norms for EQ-5D (Discussion Paper 172), 1999. https://www.york.ac.uk/che/pdf/DP172.pdf. Accessed 10 Apr 2020.

18. Kaltenthaler E, et al. Computerized cognitive behavior therapy for depression and anxiety update: a systematic review and economic evaluation. Health Technol Assess. 2006;10(33):1–70.

19. Office for National Statistics. Consumer price inflation, UK: May 2015, 2015. https://www.ons.gov.uk/economy/inflationandpriceindices/bulletins/consumerpriceinflation/2015-06-16. Accessed 3 Apr 2020.

20. Office for National Statistics. Consumer price inflation, UK: May 2019, 2019. https://www.ons.gov.uk/economy/inflationandpr*iceindices/bulletins/consumerpriceinflation/may2019. Accessed 3 Apr 2020.

21. HM Treasury. THE GREEN BOOK CENTRAL GOVERNMENT GUIDANCE ON APPRAISAL AND EVALUATION 2018. 2018. www.gov.uk/government/publications. Accessed 10 Apr 2020.

22. Office for National Statistics. National life tables: UK, 2019. https://www.ons.gov.uk/peoplepopulationandcommunity/birthsdeathsandmarriages/lifeexpectancies/datasets/nationallifetablesunitedkingdomreferencetables. Accessed 3 Apr 2020.

23. Michal M, et al. Symptoms of depression and anxiety predict mortality in patients undergoing oral anticoagulation: results from the thrombEVAL study program. Int J Cardiol. 2015;187(1):614–9.

24. Pharmaceutical Services Negotiating Committee. Professional Fees (Drug Tariff Part IIIA), 2019. https://psnc.org.uk/dispensing-supply/endorsement/fees-allowances/. Accessed 3 Apr 2020

25. OpenPrescribing.net. EBM DataLab, University of Oxford, 2020. https://openprescribing.net/. Accessed 3 Apr 2020.

26. National Institute for Health and Care Excellence. CG 113. Generalised anxiety disorder and panic disorder in adults: management, NICE, 2011. https://www.nice.org.uk/guidance/cg113. Accessed 3 Apr 2020.

27. NHS Improvement. National Cost Collection 2018/19, 2019. https://improvement.nhs.uk/resources/national-cost-collection/. Accessed 3 Apr 2020.

28. Navarro-Haro MV, et al. Evaluation of a mindfulness-based intervention with and without virtual reality dialectical behavior therapy® mindfulness skills training for the treatment of generalized anxiety disorder in primary care: a pilot study. Front Psychol. 2019;10:55.

29. Topper M, Emmelkamp PMG, Watkins E, Ehring T. Prevention of anxiety disorders and depression by targeting excessive worry and rumination in adolescents and young adults: a randomized controlled trial. Behav Res Ther. 2017;90:123–36.

30. Jankovic D, Bojke L, Marshall D, Saramago Goncalves P, Churchill R, Melton H, Brabyn S, Gega L. Systematic review and critique of methods for economic evaluation of digital mental health interventions. Appl Health Econ Health Policy. 2021;19(1):17–27.

31. Thokala P, Ochalek J, Leech AA, Tong T. Cost-effectiveness thresholds: the past, the present and the future. Pharmacoeconomics. 2018;36(5):509–22.

32. Paulden M. Why it's time to abandon the ICER. Pharmacoeconomics. 2020;38:781–4.

33. Drummond MF, Sculpher MJ, Claxton K, Stoddart GL, Torrance GW. Methods for the economic evaluation of health care programmes. Oxford: OUP; 2015.

34. Lomas J, Martin S, Claxton K. Estimating the marginal productivity of the English National Health Service from 2003 to 2012. Value Health. 2019;22(9):995–1002.

35. Vera-Llonch M, Dukes E, Rejas J, Sofrygin O, Mychaskiw M, Oster G. Cost-effectiveness of pregabalin versus venlafaxine in the treatment of generalized anxiety disorder: findings from a Spanish perspective. Eur J Health Econ. 2010;11(1):35–44.

36. Claxton KP, Sculpher MJ. Using value of information analysis to prioritise health research: some lessons from recent UK experience. Pharmacoeconomics. 2006;24(11):1055–68.

37. Walters K, Rait G, Griffin M, Buszewicz M, Nazareth I. Recent trends in the incidence of anxiety diagnoses and symptoms in primary care. PLoS ONE. 2012;7(8): e41670.

38. Strong M, Oakley JE, Brennan A. Estimating multiparameter partial expected value of perfect information from a probabilistic sensitivity analysis sample: a nonparametric regression approach. Med Decis Mak. 2014;34(3):311–26.

39. Richards D, Richardson T. Computer-based psychological treatments for depression: a systematic review and meta-analysis. Clin Psychol Rev. 2012;32(4):329–42.

40. Baumann M, Stargardt T, Frey S. Cost-utility of internet-based cognitive behavioral therapy in unipolar depression: a Markov model simulation. Appl Health Econ Health Policy. 2020;18(4):567–78.

Authors and Affiliations

Dina Jankovic[1] · **Pedro Saramago Goncalves[1]** · **Lina Gega[2]** · **David Marshall[3]** · **Kath Wright[3]** · **Meena Hafidh[1]** · **Rachel Churchill[3]** · **Laura Bojke[1]**

Pedro Saramago Goncalves
pedro.saramago@york.ac.uk

Lina Gega
lina.gega@york.ac.uk

David Marshall
d.marshall@york.ac.uk

Kath Wright
kath.wright@york.ac.uk

Meena Hafidh
Meena.Hafidh1@myport.ac.uk

Rachel Churchill
rachel.churchill@york.ac.uk

Laura Bojke
laura.bojke@york.ac.uk

[1] Centre for Health Economics, The University of York, York YO10 5DD, UK

[2] Department of Health Sciences and Hull York Medical School, University of York, York, UK

[3] Centre for Reviews and Dissemination, University of York, York, UK