Exploring Outcomes and Preferences for Guided Self Help

Linda Davies[1]
Karina Lovell[1]
Peter Bower[1]
Michael Barkham[3]
Bonnie Sibbald[1]
David Richards[2]
Chris Roberts[1]
Anne Rogers[1]
Judith Gellatly[1]
Sue Hennessay[2]
Nagina Khan[1]

Universities of Manchester[1], University of York[2], University of Leeds[3]

CORRESPONDING AUTHOR
Linda Davies
Health Economics Research at Manchester
Health Methodology Research Group
School of Community Based Medicine
University of Manchester
Oxford Road
Manchester M13 9PL

Email: linda.davies@manchester.ac.uk

INTRODUCTION

Mental health problems such as anxiety and depression are highly prevalent in primary care settings.[1] Psychological therapies are effective[2] and popular with patients,[3] but there are significant problems relating to access. Overcoming these problems requires the development and evaluation of minimal interventions (such as guided self-help) for use in a stepped approach to care. The UK National Institute for Clinical Excellence (NICE) recommends a 5 step approach to treatment and management of depression. In the primary care setting, this includes:

- Recognition and assessment of the problem (step 1),
- Watchful waiting, guided self help, computerised cognitive behavioural (CBT) therapy, exercise and brief psychological interventions for mild depression (step 2)
- Medication, psychological interventions and social support for moderate and severe depression (step 3)
- Medication, complex psychological interventions, combined treatments or ECT for people with treatment resistant, recurrent, atypical and psychotic depression and those at significant risk of self harm or suicide or demonstrating severe self neglect. (steps 4 and 5).

Adoption of the minimal interventions outlined in step 2 is predicated on the assumption that such interventions are effective, efficient, and acceptable to patients. There is a developing evidence base concerning the clinical effectiveness and cost-efficiency of minimal interventions.[4]

A recent systematic review suggested that there is no clear evidence that psychotherapy and counselling in primary care are cost effective and methodological shortcomings of the studies included in the review limited the generalisibility of the results[5]. In contrast a number of trials and a systematic review and economic modelling study suggested that the minimal interventions and technologies used in GSH can be effective and cost effective[6-11].

Furthermore the literature on the acceptability of minimal interventions in general and guided self help in particular is also limited. Minimal interventions and conventional psychological therapies differ on a number of important characteristics which may be

related to patient acceptability. These include waiting times, the background and experience professionals, the amount of face to face contact, convenience of access, and potential stigma. Patients may also differ in their expectancies as to outcomes. Patients may be willing to trade-off certain treatment characteristics against others in order to meet their mental health needs. For example, some patients may be willing to receive a treatment which involves less contact with a health professional, if it allows more rapid access to treatment. In addition, different groups of patients may have different priorities and trade-offs.

Guided self-help is defined as a CBT-based self-help resource with limited support from a health care professional. However, there is ambiguity concerning the most appropriate method to deliver the self-help material (written materials or multimedia), the level and nature of the guidance required, and the skills and expertise required to deliver this guidance. Following the UK Medical Research Council recommendations for the phased development of interventions exploratory research to devise an evidence based guided self-help intervention was conducted (Minimal Intervention for Depression and Anxiety Symptoms, the MIDAS project)[12-13]. Phases 1 and 2 of MIDAS involved theoretical and empirical modelling to define the active ingredients of GSH and an exploratory randomised controlled trial (RCT).The latter was used to test the intervention, examine delivery in routine settings, and provide estimates of key trial parameters such as recruitment rates and estimates of effectiveness, to inform the design of a definitive trial. The trial used a short period of follow up (12 weeks from entry to the trial). Generally, adherence to the MIDAS intervention protocol was high as was acceptability to both professionals and patients; the professionals delivered the majority of the specific components, the effect size of the intervention on outcomes was small, showing a modest effect of guided self-help in primary care.

The work in phases 1 and 2 of the MIDAS project identified that feeling better at the end of treatment and improved ability to manage home and work life were both important outcomes to patients and health care professionals. However, there was no clear consensus as to which outcome or outcome measure should be used as the primary outcome for the exploratory trial. Accordingly measures of feeling better (Becks Depression Inventory) and ability to manage (SAS) were included in the exploratory trial as well as a generic health status index (EQ-5D) and associated utility tariffs.

Additionally, phases 1 and 2 indicated that aspects of the process of care could be important indicators of the acceptability and use of GSH. These included waiting time for access to treatment, length of time spent with a health care professional and years of experience of that professional. A discrete choice experiment (DCE) [14-17] was designed to explore patient priorities, trade-offs and monetary valuations.

AIMS

The aim was to inform the design of the main clinical trial about the relative importance of different outcomes and the appropriate primary clinical endpoint of the trial, and the design of a subsequent stated preference survey for the main trial. A discrete choice experiment was used to examine the priorities and trade-offs made by patients among key attributes of the outcomes and process of psychological therapy and mental health care.

Specific research questions included:

1. Does the Euroqol and associated quality adjusted life year correlate with changes in the outcomes of feeling better and ability to manage and can it detect differences between treatment and control groups?
2. What measures of outcome and processes of care are important to patients and are they likely to trade off different aspects of health improvement and processes of care?

METHODS

Quality adjusted life years (QALYs)

Health status was measured by the Euroqol questionnaire at baseline and 12 weeks. The Euroqol is a validated generic health status measure, used in national health surveys in the UK, in clinical trials in mental health, and by NICE in technology appraisals and subsequent guidance. The Euroqol covers 5 domains (mobility, self care, usual activity, pain/distress, anxiety/depression). The health status profiles were converted to utility values using the published utility tariffs for the Euroqol. The QALYs were estimated as:

$$QALY=\Sigma\ ((U_i + U_{i+1})/2)*(t_{i+1} - t_i),$$

where U = utility value and t = number of days between assessments.

Patient preferences for outcome and process

The outcome measures and process attributes for the DCE were derived from a meta-synthesis of qualitative literature and systematic review of effectiveness literature and a Delphi survey conducted in Phase 1 and 2 of the MIDAS project. These sources were used to inform the design and choice of outcome measures used for the exploratory trial. The data from these three sources were used to develop an initial list of key process and outcome attributes to be included in the discrete choice questionnaire (Appendix 1). The list was reviewed by all members of the MIDAS team and discussed at length, in terms of: relevance to the aims of MIDAS; breadth of coverage and whether there were missing attributes; appropriate levels for each attribute; ambiguity in the phrasing and wording of the descriptions of the attributes and levels and whether the attributes effectively described the issues to be captured. The study team reached consensus on the final set of attributes and levels to be included in the questionnaire.

The DCE used discrete choice questions. Participants were presented with a sequential series of questions. Each question included a pair of profiles, A and B. Each profile in the pair described a hypothetical health state in terms of the outcomes and therapy processes. All of the profiles in the survey used the same attributes. The levels for the attributes differed between the pairs of profiles in each question, and between the questions. Respondents were asked to choose their preferred option (A or B) from the two profiles. The data from the survey provides information about the extent to which respondents were willing to trade between positive and negative outcomes for each of the attributes.

A fractional factorial design was designed to meet 3 criteria: (1) near orthogonal arrays of discrete choice sets, subject to the constraints of excluding implausible attributes and levels or uninformative combinations of pairs (2) balance in the number of times a level appears (3) minimal overlap between the levels in each choice. The orthogonal array of attributes and levels and number of choice sets were determined using published design catalogues (http://www.research.att.com/~njas/oadir/) and a fold-over design. Demographic data about age, ethnicity and whether the whether the respondent had previously ever had depression or anxiety and whether they had been treated for anxiety or depression was also collected. An example of the DCE questions is shown in Appendix 2. The DCE questions were piloted with a convenience sample of

respondents to an electronic survey sent to 11,330 staff and students in the University of Manchester, School of Nursing. The data from the pilot study were analysed to assess the consistency and reliability of the responses, the acceptability of the choice sets included in the questionnaire as measured by the number of invalid, indifferent or non responses to each choice set. The written qualitative comments made by the pilot survey participants were also reviewed. No changes were made to the DCE questions following the analysis of the pilot data.

The participants in the exploratory trial were asked to complete the final DCE questionnaire at baseline and 3 month follow up. Only the baseline data are reported here.

Clinical opinion was used to define the range of BDI and SAS scores for each of the levels included in the DCE for the attributes of feeling better and ability to manage. The willingness to pay values for each level were used to weight and combine the BDI and SAS scores into a single monetary measure (M-BDI-SAS). The joint outcome measure was used to explore whether a measure that combined feeling better and ability to manage would identify statistically significant differences between the control and GSH groups in the exploratory trial.

Analysis

Descriptive statistics and regression were used to explore whether the Euroqol and QALYs were associated with the outcomes of feeling better and ability to manage home and work life and the single monetary measure, M-BDI-SAS. Analysis of covariance (ANCOVA) was used to evaluate whether QALYs detected differences between the control and GSH groups at 3 months follow up.

There were no changes to the DCE questions between the pilot and exploratory trial surveys, so the data from each were pooled for the main analysis. Descriptive statistics were used to summarise the demographic characteristics of the pilot and trial samples at baseline. Multivariate analysis (MANOVA) of attributes rank scores was used to assess whether the rank of attributes varied by participants characteristics (age, respondent type, gender, ethnic group, ever had or ever treated for depression). Since this was an exploratory rather than explanatory analysis, a high cut off value for $p <= 0.10$ was used. Logistic regression was used to assess whether the DCE choices made varied by

demographic characteristics, respondent type or whether the respondent had ever had or been treated for depression.

A comparison of binomial logistic regression and probit models gave similar marginal effects on choice (logit y = 0.451, probit y = 0.450). The Hausman test was used to compare fixed and random effects logistic regression models and indicated that a random effects model was not likely to be consistent. Therefore the discrete choice data were analysed using a conditional logistic regression model. The analysis included estimation of the attribute and level coefficients for the main effects. Interactions were explored if there was evidence that choices were influenced by respondent characteristics. The analysis also explored the issue of whether the attributes of feeling better and ability to manage were independent. The exact form of the models was determined by the results of the pilot study and descriptive analysis of the full survey data. The monetary values for each attribute were estimated as the marginal rate of substitution between that attribute and the cost attribute.

Analysis of covariance (ANCOVA) was used to evaluate whether the single monetary measure (M-BDI-SAS) detected differences between the control and GSH groups at 3 months follow up.

RESULTS

Health status, utility and QALYs in the exploratory trial

At baseline, utility scores were statistically significantly correlated with the BDI and SAS ($p<0.01$). The baseline utility measures were also correlated with other measures of outcome (CORE, PHQ, $p<0.05$), and the main employment category of the participant ($p<0.01$). There were no statistically significant correlations between the follow up utility scores and clinical measures of depression and anxiety. QALYs between baseline and follow up were correlated with the BDI ($p<0.05$). Utility scores at baseline ($p<0.06$) and follow up ($p=0.000$) and QALYs ($p<0.05$) were associated with the M-BDI-SAS. Table 5 gives the unadjusted utility scores and QALYs for the complete case data.

A generalised linear model was used to control for the baseline clinical and patient characteristics that were statistically significantly correlated ($p<0.05$) with QALYs. The

analysis was also controlled for baseline utility and length of follow up. The GSH group had higher baseline utility and longer duration of follow up than the control group (although the proportion of patients followed up was similar in both groups). Adjusting for baseline demographic and clinical variables as well as baseline utility then GSH was associated with a statistically significant higher mean QALY than the control group (mean difference = 0.043, 95% CI = 0.034 to 0.053, p=0.000). However, there was no statistically significant difference between the GSH and control group on the primary and secondary clinical outcome measures, including the BDI and SAS.

Table 1 Utility and QALYS by allocation group, unadjusted data

	Baseline Utility		Follow up utility		QALYs	
	Control	GSH	Control	GSH	Control	GSH
Complete case						
Mean	0.62	0.66	0.66	0.67	0.17	0.21
Std. Deviation	0.32	0.24	0.31	0.28	0.09	0.07
N	28	28	20	18	19	17

Patient preferences for process and outcomes

Table 2 shows the demographic characteristics of the survey participants. The response rate was low for the pilot survey (104/11330) which is not surprising for an unsolicited email invitation of this nature. A total of 58 patients were recruited to the exploratory trial. Of these 47 (81%) completed the ranking exercise, 48 (83%) answered at least one binary choice question and 42 (72%) completed all the binary choice questions. The binary DCE choices were equally distributed between option A and option B, overall and by participant type (patient or pilot sample; Pearson chi square; p=0.985).

Table 2 Characteristics of respondents to stated preference survey

Characteristic	Patients		Pilot		Total		P=
	No. (%)	n=	No. (%)	n=	No. (%)	n=	
Female	43 (73)	58	16 (27)	104	59 (36)	165	0.000
White	54 (93)	58	99 (95)	104	153 (93)	165	0.069
Ever had depression	58 (100%)	58	47 (44)	104	105 (64)	165	0.000
Ever treated	58 (100%)	58	48 (45)	104	106 (64)	165	0.000
	Mean (sd)	**n=**	**Mean (sd)**	**n=**			
Age	37 (12)	58	32 (20)	78	34(17)	136	0.050

Table 3 summarises how the participants ranked the attributes included in the survey. There were no statistically significant differences by age, gender or ethnic group ($p>0.10$). There were statistically significant differences ($p<0.10$) between the patient and pilot samples in the ranking of whether a treatment led to feeling better, time spent with the health professional and cost. However, there were no statistically significant differences between participants in the binary DCE choices made (logistic regression, $p>0.30$ for all demographic characteristics, respondent type and ever had or ever treated for depression). On the basis of the rank data, an initial model including interactions with respondent type was specified. Preferences were related to respondent type (LR $chi^2(12) = 26.15$; $p= 0.0102$) for the attributes of waiting time, cost, feeling better and ability to manage ($p<0.001$).

The attributes of feeling better and ability to manage were included in the DCE because there was uncertainty about which was the more relevant measure to use as the primary outcome in a clinical trial. Some of the qualitative comments from the focus groups before the feasibility study and open text comments on the DCE suggested that these two attributes may be alternative measures of the same outcome or concept, whereas others suggested that they were linked but different outcomes. The analysis indicated that the coefficients on both attributes are statistically significant. Comparison of the full model including both attributes versus a constrained model excluding ability suggested that both attributes are important and should be included (LR $chi^2(2) = 1332$; $p= 0.000$).

Table 3 Rank of attributes, by type of respondent

	Rank (number, %)[1]						
	p=[2]	**1**	**2**	**3**	**4**	**5**	**6**
Time to get access to treatment							
Patients	0.37	6(13)	5(11)	14(30)	14(30)	7(15)	1(2)
Pilot		11(10)	14(13)	45(42)	22(21)	10(9)	4(4)
All		17(11)	19(12)	59(39)	36(24)	17(11)	5(3)
Cost to you of treatment							
Patients	0.08	4(9)	1(2)	10(21)	11(23)	7(15)	14(30)
Pilot		11(10)	8(8)	25(24)	35(33)	16(15)	11(10)
All		15(10)	9(6)	35(23)	46(30)	23(15)	25(16)
Total time spent with professional during treatment							
Patients	0.00	5(11)	3(6)	10(21)	7(15)	15(32	7(15)
Pilot		1(1)	6(6)	9(8)	25(24)	47(44)	18(17)
All		6(4)	9(6)	19(12)	32(21)	62(41)	25(16)
How much better you feel at the end of treatment							
Patients	0.07	26(55)	17(36)	2(4)	1(2)	0(0)	1(2)
Pilot		37(35)	40(38)	16(15)	7(7)	6(6)	0(0)
All		63(41)	57(37)	18(12)	8(5)	6(4)	1(1)
Your ability to manage home & work life at the end of treatment							
Patients	0.36	18(38)	17(36)	6(13)	2(4)	2(4)	2(4)
Pilot		48(45)	36(34	10(9)	7(7)	2(2)	3(3)
All		66(43)	53(35)	16(10)	9(6)	4(3)	5(3)
Years experience of health professional in managing depression							
Patients	0.00	1(2)	2(4)	3(6)	8(17)	15(32)	18(38)
Pilot		2(2)	0(0)	2(2)	7(7)	25(24)	70(66)
All		3(2)	2(1)	5(3)	15(10)	40(26)	88(58)

Note to Table 3

1. 1 = most important, 6 = least important
2. Multivariate analysis of participant characteristics of age, respondent type, gender, ethnic group

Table 4 shows the main effects from the regression model of attributes on the DCE binary choice. The sign and magnitude of the coefficients indicates that participants preferred:

- Lower waiting times and costs;
- Feeling better;
- Being better able to manage;
- Spending more time with the health professional, although lower the coefficient for spending 9 hours rather than 3 hours with the professional, suggests the marginal benefits of time spent decrease as the number of hours increases;
- The health professional to have more than 2 years experience managing depression, although the lower coefficient for the health professional having 10 years experience, suggests the marginal benefits of experience decreases as the years of experience increases.

Table 4 Conditional logit model, no interactions, all respondents

Choice		Coeff.	s.e.	P>|z|	95% CI	
Ability to manage[1]	A little better	1.561	0.096	0.00	1.374	1.748
	A lot better	3.058	0.100	0.00	2.862	3.254
Feeling better[1]	A little better	1.058	0.097	0.00	0.869	1.248
	A lot better	2.255	0.097	0.00	2.065	2.445
Time to access treatment [2]	3 months	-0.313	0.085	0.00	-0.479	-0.147
	6 months	-0.503	0.093	0.00	-0.685	-0.321
Experience of health professional[3]	4 yrs	0.401	0.097	0.00	0.212	0.591
	10 yrs	0.078	0.095	0.41	-0.108	0.264
Time spent with professional[4]	3 hours	0.573	0.097	0.00	0.382	0.763
	9 hours	0.249	0.094	0.01	0.065	0.434
Cost to you of treatment [5]	£500	-0.723	0.085	0.00	-0.889	-0.557
	£1000	-1.429	0.097	0.00	-1.620	-1.239

LR $chi^2(12)$ = 2300.23, Prob > chi^2 = 0.0000, Pseudo R^2 = 0.3436

Notes to Table 4
1. Reference case for feeling better at end of treatment and ability to manage is no improvement
2. Reference case for waiting time to access treatment is 1 month
3. Reference case for years experience of health professional is 2 years
4. Reference case for time spent with health professional is 1 hour
5. Reference case for cost to you of treatment is £0

Table 5 shows the estimated willingness to pay and willingness to wait for different attributes of process and outcome.

Table 5 Willingness to pay preferred levels of attributes

Choice	Willingness to pay, £'s		
	Patient	**Pilot**	**All**
Your ability to manage home & work life at the end of treatment			
Improve to a little better	1032	1036	1042
Improve to a lot better	1877	2097	2041
How much better you feel at the end of treatment			
Improve to a little better	795	667	706
Improve to a lot better	1652	1426	1505
Time to get access to treatment			
Increase to 3 months	-83	-256	-209
Increase to 9 months	-216	-380	-336
Years experience of health professional in managing depression			
Increase yrs experience to 4	479	254	268
Increase yrs experience to 10	289	-16	52
Total time spent with professional during treatment			
Increase time spent to 3 hours	479	342	382
Increase time spent to 9 hours	289	115	166

The single monetary measure, M-BDI-SAS, was used to explore whether a measure that combined feeling better and ability to manage would identify statistically significant differences between the control and GSH groups in the exploratory trial. Table 6 shows the results of an ANCOVA to control for differences in baseline covariates. The results suggest that GSH was associated with improvements in outcome using the combined WTP values for feeling better and ability to manage ($p<0.01$).

Table 6 Total of BDI and SAS weighted by WTP values for feeling better and ability to manage

Combined WTP values for BDI & SAS	Coefficient	s.e.	P>z	95% CI
Days follow up	-65	22	0.00	-109; -22
Age	-56	74	0.45	-200; 89
Gender	-2611	876	0.00	-4329; -894
Time on antidepressants 2-4 weeks	5182	1300	0.00	2635; 7730
Time on antidepressants 1-3 months	2710	1046	0.01	660; 4760
Time on antidepressants 4-6 months	1767	1020	0.08	-232; 3767
Time on antidepressants 7-12 months	3011	1181	0.01	695; 5326
Baseline BDI	-10	81	0.90	-169; 149
Baseline SAS	-103	1001	0.92	-2064; 1858
Baseline utility	1034	2585	0.69	-4032; 6100
Allocation group=GSH	2170	843	0.01	517; 3823
Constant	8961	11672	0.44	-13916; 31837

CONCLUSIONS

Do the Euroqol and QALY correlate with clinical outcomes and detect differences between treatment and control groups?

The data indicated that there may be differences in the health status and related utility of patients receiving guided self help compared to usual care at baseline and follow up. The Euroqol and associated population utility values are validated measures for the estimation of QALY's and are used in evaluations of depression. However, as generic measures they may not be sensitive to small but important changes in health and health related quality of life. There were small differences in QALY's between the GSH and control groups. Analysis of the trial data indicated that the utility values correlated with other outcome measures at baseline and that changes in health status and in utility were detected by the EQ-5D over the 3 month follow-up period. The QALY measure also detected statistically significant differences between the GSH and control groups. These factors support the use of the Euroqol and associated utility tariff to measure QALYs. However, there were no statistically significant differences in the primary and secondary

clinical outcome measures used in the trial. One criticism of the Euroqol is that it is restricted to 3 levels of effect on each domain (ie no problems, some problems, extreme problems). For people who are at the top or bottom of each level, small changes in health status mean that the person is forced into the next level, giving an artificial inflation in the size of any change. Alternatively, the Euroqol may have captured a range of small but important changes in health on different domains. Each of these changes taken individually may not be sufficient to identify a statistically significant or clinically important effect on each of the clinical outcome measures. However, taken together in a single measure of utility or QALY, the full impact of these small changes is identified. This is supported by the results of the DCE.

What measures of outcome and process of care are important?

Both the pilot and patient samples appeared able to prioritise and trade between the different measures of outcome. The DCE found that feeling better and ability to work and engage in usual activities were both important measures of the outcome of treatment. Overall, the attribute coefficients and WTP values suggest that ability to manage may be slightly more important than feeling better at the end of treatment. However, the data may also suggest that both outcomes should be considered in the evaluation of the effectiveness of treatment. As with the utility and QALY measures, the combined monetary value of improvements in feeling better and ability to manage indicated a statistically significant difference between the control and GSH groups, whereas there was a trend but no statistically significant difference in the separate clinical outcomes of the BDI and SAS. This may suggest that the primary outcome of interest for future evaluations and assessment of the effectiveness of treatment is a measure that combines these two types of outcome.

In addition, the process measures of waiting time for treatment, time spent with a health care professional and years of experience of a health professional were statistically significant and important attributes. However, there appeared to be declining marginal benefit for the latter two attributes.

There appeared to be differences between the pilot and patient groups in the ranking given to these two outcome measures of feeling better and ability to manage at the end

of treatment. These were reflected to some extent in the differences between WTP values for different levels of these outcomes. There were also differences between the samples in WTP for reduced waiting times, suggesting that patients were more willing to wait for treatment than the pilot sample.

The differences in preferences between the two samples may have important policy implications in terms of what is considered an important outcome and acceptable waiting time for treatment. The pilot sample comprised academic staff and students in a University School of Nursing, rather than patients, if the pilot sample preferences and valuations are representative of health care professionals who would provide interventions such as GSH. It is well documented that patients and non patients often have different valuations of the same health states and the differences in preferences may represent differences between societal and patient preferences.

ISSUES

There are a number of limitations to the data and the analyses which include amongst others

- The small sample of participants enrolled in the exploratory clinical trial, for whom clinical and utility data are available; There were also differences between the control and GSH groups in baseline utility and length of follow up. These two factors limit the robustness of the analyses of the utility and QALY data.
- The small sample of participants in the DCE survey. This combined with the use of pooled pilot and patient data and relatively small sample of patients mean that the data and analyses can only be considered as exploratory and be used as pointers for further research.
- The inclusion of the cost attribute in the DCE.
- The use of the conditional logistic regression model.

Given these limitations is it worth pursuing further analyses of the current data and further data collection to explore:

- The relative pros and cons of using the QALY versus single clinical outcome measures as the primary outcome for future trials in mild to severe depression.

- The need for a new preference based measure that measures and combines the attributes of feeling better and ability to manage.
- The use of more complex, but potentially more informative regression models to explore preference heterogeneity between participants in the DCE survey.
- Further qualitative and DCE data collection to explore
    - whether there are differences between patients and societal or professional preferences and valuations and the implications of these for health policy and clinical decision making.
    - The validity and interpretation of the cost attribute in this setting.

References

1. Goldberg D, Huxley P. *Common Mental Disorders: A biosocial model*. London; Routledge, 1992.

2. Department of Health. *Treatment Choice in Psychological Therapies and Counselling: Evidence based clinical practice guideline.* London. Department of Health, 2001.

3. Priest R, Vize C, Roberts A, Roberts M, Tylee A. Lay people's attitudes to treatment of depression: results of opinion poll for Defeat Depression Campaign just before its launch. *BMJ* 1996;**313**:858-859.

4. Anderson L, Lewis G, Araya R, et al. Self-help books for depression: how can practitioners and patients make the right choice? *Br J Gen Pract* 2005;**55**:387-392.

5. Bosmans, Judith E., et al. Are Psychological Treatments for Depression in Primary Care Cost Effective? [Journal Article] Journal of Mental Health Policy and Economics. Vol. 11 (1). p 3-15. March 2008

6. W J. Katon; P Roy-Byrne; J Russo; DCowley, Cost-effectiveness and Cost Offset of a Collaborative Care Intervention for Primary Care Patients With Panic Disorder. Arch Gen Psychiatry. 2002;59:1098-1104

7. P McCrone, M Knapp, J Proudfoot, C Ryden, K Cavanagh, D A. Shapiro, S Ilson, J A. Gray, D Goldberg, A Mann, I Marks, B Everitt A Tylee. Cost-effectiveness of computerised cognitive behavioural therapy for anxiety and depression in primary care: randomised controlled trial. B J Psychiatry (2004), 185, 55-62

8. L Hakkaart-Van Roijen, A Mieke Van Straten, M Al, F Rutten, M Donker Cost-utility of brief psychological treatment for depression and anxiety. B J Psychiatry (2006), 18 8 , 323-329

9. F Smit, G Willemse, M Koopmanschap, S Onrust, P Cuipers, A Beekman Cost-effectiveness of preventing depression in primary care patients. B J Psychiatry (2006), 188 , 330-336

10. Kaltenthaler E. Brazier J. De Nigris E. Tumur I. Ferriter M. Beverley C. Parry G. Rooney G. Sutcliffe P. Computerised cognitive behaviour therapy for depression and anxiety update: a systematic review and economic evaluation. [Journal Article, Glossary, Research, Systematic Review, Tables/Charts] Health Technology Assessment. 2006; 10(33): iii, xi-xiv, 1-181

11. K Lovell, P Bower, D Richards, M Barkham, B Sibbald, C Roberts, L Davies, A Rogers, J Gellatly, S Hennessy Developing guided self-help for depression using the Medical Research Council complex interventions framework: a description of the modelling phase and results of an exploratory randomised controlled trial BMC Psychiatry 2008, 8:91

12. Gellatly J, Bower P, Hennessey S, Richards D, Gilbody S, Lovell K: What makes self-help interventions effective in the management of depressive symptoms? Meta-analysis and meta-regression. *Psychol Med* 2007, 37:1217-1228.

13. Khan N, Bower P, Rogers A: Guided self-help in primary care mental health: a meta synthesis of qualitative studies of patient experience. *Br J Psychiatry* 2007, 191:206-211.

14. Ryan M, Bate A, Eastmond C, Ludbrook A. Use of discrete choice experiments to elicit preferences. *Quality in Health Care* 2001;**10**:i55-i60

15. Ryan M, Farrar S. Using conjoint analysis to elicit preferences for health care. *BMJ* 2000;**320**:1530-1533.

16. Coast J, Horrocks S. *Developing attributes and levels for discrete choice experiments: a case study using qualitative methods.* Department of Social Medicine. University of Bristol, 2005.

17. Carlsson F, Martinsson P. Design techniques for stated preference methods in health economics. *Health Econ* 2004;**12**:281-294.

Appendix 1 Possible attributes and levels derived from phase 1 and 2 of MIDAS

Attribute	Possible wording	Possible levels
Access	You get psychological therapy treatment in….	2 weeks 1 month 3 months 6 months 12 months
Session duration	Your sessions with the [mental health professional] last Y minutes	15 minutes 30 minutes 45 minutes 60 minutes
Session number	You have X sessions with the [mental health professional]	1 3 6 12
Session duration and number	You have x sessions with the [mental health professional] which last y minutes	See above
Professional experience	Your treatment is provided by a [mental health professional] with…	Limited experience A lot of experience 2 years experience 10 years experience
Willingness to pay	The treatment costs…	£0 £8 £18 £28
Clinical outcome	At 3 months you feel	Worse No different Quite a bit better than you were A lot better than you were Not depressed Mildly depressed Moderately depressed Severely depressed
Satisfaction	At 3 months you feel	Satisfied with the treatment Unsatisfied with the treatment Satisfied with the way you were dealt with Unsatisfied with the way you were dealt with
Social function	At 3 months you feel	No change in your ability to manage work and home Slight change in your ability to manage work and home Major change in your ability to manage work and home
Therapeutic alliance	The therapist is	Caring and sensitive

Appendix 2 Example DCE question

Consider the following characteristics that describe 2 options for treatment of depression. Which of these options do you prefer, OPTION A or OPTION B? Please tick (√) in the box for the option you like the most.

	OPTION A	OPTION B
Time to get access to treatment	In 1 month	In 3 months
Cost to you of treatment	£0	£500
Total time spent with professional during treatment	1 hour	3 hours
How much better you feel at the end of treatment	No better	A little better
Your ability to manage your home and work life at the end of treatment	No change	A little better able to manage
The number of years experience of your health professional in dealing with depression	2 years experience	4 years experience
Tick (✓) ONE box only		