ORIGINAL RESEARCH ARTICLE

Cost-Utility Analysis of Discontinuing Antidepressants in England Primary Care Patients Compared with Long-Term Maintenance: The ANTLER Study

Caroline S. Clarke[1] · Larisa Duffy[2] · Glyn Lewis[2] · Nick Freemantle[3] · Simon Gilbody[4] · Tony Kendrick[5] · David Kessler[6] · Michael King[2] · Paul Lanham[7] · Derelie Mangin[8] · Michael Moore[5] · Irwin Nazareth[1] · Nicola Wiles[9] · Louise Marston[1] · Rachael Maree Hunter[1]

Abstract

Background Depression is a common mental health condition with considerable negative impact on health and well-being. Although antidepressants are recommended as first-line treatment, there is limited evidence regarding the cost effectiveness of long-term maintenance antidepressants for preventing relapse.

Objectives Our objective was to calculate the mean incremental costs and quality-adjusted life-years (QALYs) over 12 months of discontinuing long-term antidepressant medication in well patients compared with maintenance, using patient-level trial data.

Methods We conducted a cost-utility analysis of 478 participants from 150 UK general practices recruited to a randomised, double-blind trial (ANTLER). QALYs were calculated from EQ-5D-5L and 12-Item Short Form survey (SF-12) results, with primary analysis using the EQ-5D-5L value set for England. Resource use was collected from primary care patient electronic medical records and self-completed questionnaires capturing mental-health-related resource use. Costs were calculated by applying standard UK unit costs to resource use. Adjustments were made for baseline variables.

Results Participants randomised to discontinuation had significantly worse utility scores at 3 months (− 0.032; 95% confidence interval [CI] − 0.053 to − 0.011) but no significant difference in QALYs (− 0.011; 95% CI − 0.026 to 0.003) or costs (£3.11; 95% CI − 41.28 to 47.50) at 12 months. The probability that discontinuation was cost effective compared with maintenance was 12.9% at a threshold of £20,000 per QALY gained.

Conclusions Discontinuation of antidepressants was unlikely to be cost effective compared with maintenance for currently well patients on long-term antidepressants. However, this analysis provides no information on the wider impact of antidepressants. Our findings provide information on the potential impact of discontinuing long-term maintenance antidepressants and facilitate improving guidance for shared patient–clinician decision making.

Trial Registration EudraCT number 2015-004210-26; ISRCTN number ISRCTN15969819.

✉ Caroline S. Clarke
caroline.clarke@ucl.ac.uk

Extended author information available on the last page of the article

Key Points for Decision Makers

There was no evidence of cost savings for the English National Health Service on discontinuing treatment, and there was a potential disadvantage to patients' health-related quality of life in the short term.

Discontinuation of maintenance antidepressants for currently well patients is unlikely to be recommended nationally on the grounds of cost effectiveness. Our findings can inform guidance on discontinuation of maintenance antidepressants, with the aim of facilitating joint patient–clinician decision making regarding maintenance antidepressants, alongside other considerations, such as potential longer-term adverse effects of medication, that could influence the decision that were not captured in this study.

1 Introduction

Depression is a common mental health condition that has considerable negative impact on the health and well-being of individuals as well as substantial negative social and financial impacts on the wider community [1]. In the UK in 2007, it cost £1.7 billion in healthcare service use costs and £7.5 billion in lost employment, projected to rise to £3 billion and £12.2 billion, respectively, by 2026 [2].

Depression is generally managed in primary care with antidepressants as part of first-line treatment alongside psychological therapies. Numbers of prescriptions for antidepressants in England are growing [3], partially because of their increased use as maintenance therapy to prevent relapses [4]. Selective serotonin reuptake inhibitors are the most commonly used and recommended antidepressants and represent a relatively small mean purchase cost of around 4 pence per day. The majority of analyses evaluating the cost effectiveness of prescribing antidepressants for depression have conducted head-to-head decision modelling of different antidepressants to determine the most cost-effective antidepressant to treat current symptoms, rather than considering the cost effectiveness of their long-term use [4–6], with analyses rarely going beyond a 12-month time horizon [7]. The impact of side effects and withdrawal symptoms following long-term use is rarely considered, so only limited, poor-quality data are available on which decision modelling to describe long-term use could be based [8].

The ANTLER study [9, 10] was a double-blind randomised controlled trial evaluating tapering of participants' antidepressant medications down to zero dose (discontinuation arm) compared with antidepressant maintenance therapy continuing with participants' current prescriptions (maintenance arm). The ANTLER study recruited patients who were currently taking one of four common antidepressants at standard doses and were well enough to consider stopping medication. Trial participants were taking oral citalopram 20 mg/day, sertraline 100 mg/day, fluoxetine 20 mg/day or mirtazapine 30 mg/day for at least 9 months before being recruited into the trial and were randomised to either maintain this treatment or to taper their dose to zero over 1 or 2 months, for replacement by an identical-looking placebo.

This paper reports the results of a trial-based cost-utility analysis (CUA) evaluating antidepressant discontinuation compared with maintenance in primary care in England over 12 months, using patient-level data on healthcare resource use and a preference-based measure of health-related quality of life (EQ-5D-5L). The main clinical results are reported separately [10] with a primary clinical outcome of time to depression relapse.

2 Methods

2.1 Trial Design and Population

Participants provided written informed consent and were recruited via 150 primary care practices in Bristol, London, Southampton and York to a double-blind, individually 1:1 randomised controlled trial, minimising on three pre-specified variables. Minimisation is a randomisation method that allocates participants to their randomised group according to prognostic factors, aiming at achieving balance across these factors [11]. It was conducted by Sealed Envelope, using site, antidepressant medication, and the median of the baseline Revised Clinical Interview Schedule (CIS-R) score, which was used as a measure of depressive symptoms. Potentially eligible participants were found via searches of electronic primary care health records and invited by their general practitioners (GPs) to take part. To be eligible, patients aged ≥ 18 years had to have been prescribed and adhered to taking antidepressants (citalopram, fluoxetine, sertraline or mirtazapine) for at least 9 months but be feeling well enough to discontinue medication. Exclusion criteria included age ≥ 75 years; other depressive, psychotic or organic mental illnesses, including meeting *International Classification of Diseases, Tenth Revision,* ICD-10 criteria for depression; contraindications to the medication or placebo ingredients; pregnancy; and breast feeding. Further details regarding randomisation, patient screening, eligibility and recruitment can be found in the protocol paper [9].

Participants who were randomised to the discontinuation arm underwent tapering of their medication over 2 months (citalopram, sertraline and mirtazapine) or 1 month (fluoxetine), with their medications replaced by matched placebo capsules to maintain blinding. Participants randomised to the maintenance arm continued to take their antidepressant medication at the same daily doses as before.

Ethical approval was obtained from the National Research Ethics Service committee, East of England, Cambridge South (ref. 16/EE/0032). Clinical trial authorisation was given by the Medicines and Healthcare products Regulatory Agency. The trial sponsor was University College London, UK. The trial was registered: EudraCT number 2015-004210-26; protocol number 14/0647 (version 7.0); Controlled Trials ISRCTN Registry, ISRCTN15969819.

2.2 Resource Use, Costs and Utilities

Resource use information was collected from primary care electronic medical records for primary care contacts and prescriptions. This covered the 12 months of the study plus 6 months preceding baseline to provide baseline costs for adjustment. The costs of the four ANTLER medications in each arm were calculated according to doses given in the protocol and according to prescription dates and other information collected from participants' primary care electronic medical records. During the 12 months of the study, ANTLER medication in the discontinuation arm was costed as citalopram, sertraline and mirtazapine for 1 month at half the original dose, followed by 1 month at a quarter of the original dose, followed by no cost for the remaining 10 months of the study (i.e. placebo administered during the trial was priced at zero for this CUA); and for fluoxetine as 1 month at half the original dose followed by no cost for the remaining 11 months of the study, unless participants in the discontinuation arm reported stopping their study tablets before the end of month 2 (or month 1 for those initially on fluoxetine). ANTLER medication in the maintenance arm was calculated as the continuation of medication at the dose prescribed at recruitment for the 12 months of the study or until the date on which participants reported stopping their medication. Use of other relevant antidepressant medications (citalopram, fluoxetine, mirtazapine, sertraline and amitriptyline, diazepam, lorazepam and zopiclone) prescribed in either arm at any point during the study was captured from participants' electronic medical records and costed according to reported daily doses and prescription information. Unit costs for medications were obtained from the British National Formulary [12] and were applied using the lowest package cost to the National Health Service (NHS), according to the duration, dose and frequency of each reported prescription.

Participants completed a modified version of the Client Service Receipt Inventory (CSRI) [13] at baseline and at 6 and 12 months, asking about resource use for the preceding 6 months at each time point for any resource use related to their mental health. The CSRI captured information on community and acute care health service contacts, mental health community and inpatient service use, social care, employment and welfare payments, covering information that could not be obtained from primary care electronic medical records.

Unit costs of healthcare contacts were obtained from the Personal Social Services Research Unit (PSSRU) [14] and NHS reference costs [15] (see Table 1). Private healthcare resource use was costed based on participants' reported out-of-pocket costs. For the very few participants who reported using private healthcare but did not report actual out-of-pocket costs, we assumed the equivalent PSSRU and NHS reference costs. Productivity was costed for a secondary analysis using the human capital approach to cost time off work with mean costs of Office for National Statistics employment categories applied according to the occupation described in free text in the CSRI [16]. All costs are in UK £, year 2018/2019 values.

Participants completed the EQ-5D-5L [17] and 12-Item Short Form survey (SF-12) [18, 19] at baseline and 3, 6, 9 and 12 months. For the primary analysis, utility scores to calculate quality-adjusted life-years (QALYs) were calculated from participants' responses to the EQ-5D-5L using the Devlin et al. [20, 21] time trade-off tariff, hereafter called the value set for England (VSE). The van Hout et al. [22] crosswalk algorithm for generating utilities from EQ-5D-5L via the EQ-5D-3L tariff is currently preferred by England's National Institute for Health and Care Excellence (NICE) so was used in a secondary analysis [22]. Participants' responses to the SF-12 were used in another secondary analysis to calculate utilities and QALYs using the SF-6D utility scoring tariff to further test the robustness of the results to choice of utility estimation method [19]. Although NICE currently recommends the van Hout et al. [22] crosswalk tariff for calculating QALYs [23], there is concern that the crosswalk algorithm is not as sensitive to changes in depression as the VSE, hence the use of the VSE as the primary analysis [24, 25].

2.3 Statistical Analysis

Analyses were pre-specified in a health economics analysis plan [25]. The analysis deviated from this plan by swapping over the primary and secondary analyses in June 2020 so that the EQ-5D-5L/VSE was the primary analysis instead of the EQ-5D-5L/crosswalk. This was because the VSE had recently been shown to have greater sensitivity than the crosswalk in mental health [24]. We considered that this

Table 1 Unit costs used in analysis

Category	Unit costs	Relevant assumptions	Source
GP surgery consultation	28.00	9 min	PSSRU 2018–19
GP phone consultation	15.50	5 min	PSSRU 2018–19
GP home visit	34.72	11.2 min (PSSRU 2015)	PSSRU 2018–19
Practice nurse surgery consultation	12.30	20 min	PSSRU 2018–19
Practice nurse phone consultation	6.17	10 min	PSSRU 2018–19
Practice nurse home visit	21.60	35 min	PSSRU 2018–19
Phlebotomist	4.00	10 min	PSSRU 2018–19
CBT therapist	54.50	Band 7	PSSRU 2018–19
CBT therapist, privately funded by patient	50.35	Mean from study data	NA (or PSSRU 2018–19 used if cost missing)
Clinical psychologist	64.68	Band 8a	PSSRU 2018–19
Exercise or physical activity scheme or "Exercise on prescription"—NHS	10.28	Uplifted to 2018–19 prices using HCHS indices from PSSRU 2018–19	Isaacs et al. [26]
NHS walk-in centres	35.38	Estimated using PSSRU	PSSRU 2018–19
Ambulance or hospital transport	257.34	NA	NHS reference costs 2018–19
NHS direct or "call 111"	13.26	Uplifted to 2018–19 prices using HCHS indices from PSSRU 2018–19	Pope et al. [27]
A&E attendance	155.70	Weighted mean of top two non-admitted categories	NHS reference costs 2018–19
Hospital admission	1909.49	Weighted mean of EL, NEL, NES, RP costs	NHS reference costs 2018–19
Mental health nurse (or CPN)	33.83	Band 5, community-based scientific and professional staff	PSSRU 2018–19
Occupational therapist	44.16	Band 8a, community-based scientific and professional staff	PSSRU 2018–19
Social worker	44.55	Social worker, adult services	PSSRU 2018–19
Other medical professional (mostly consultant-level NHS)	100.82	Mean cost from roles given in free-text descriptions from participants	PSSRU 2018–19
Other medical professional, privately funded by patient	32.67	Mean from study data	NA

Costs are presented in £, year 2018/2019 values

A&E accident and emergency, *CBT* cognitive behavioural therapy, *CPN* community psychiatric nurse, *EL* elective inpatients, *GP* general practice or general practitioner, *HCHS* Hospital and Community Health Services, *NA* not applicable, *NEL* non-elective inpatients, *NES* non-elective short stay, *NHS* National Health Service, *PSSRU* Personal Social Services Research Unit, *RP* regular day or night admissions

was an acceptable change as the NICE crosswalk algorithm is also subject to criticism and is likely to be changed again soon [28]. All analyses were based on intention to treat and corresponded with the analyses in the clinical-effectiveness paper [10]. QALYs were calculated as the area under the curve using the standard methodology set out in Hunter et al. [29]. Costs in the primary analysis were from an England health and social care cost perspective, and participants were asked to focus on services used because of their mental health issues [30]. The exception to this was that all primary care consultations were captured because it was not possible to separate out consultations that were not related to the participant's mental health. This is the standard disease-specific cost perspective used by NICE, and we used this instead of asking for all healthcare service use because we were interested in the incremental costs, i.e. the difference between the arms. We felt that unrelated resource use would

not differ substantially between the arms as no mechanism was identified for how treatment would impact on general medical costs, and would not be large compared with services used for mental health reasons, and therefore would not be relevant, thus constituting an unnecessary burden on participants. Patients could decide what they felt was related to their mental health issues and what was not. As the time horizon for the analysis was 12 months, costs and QALYs were not discounted.

Descriptive statistics are reported for adjusted, multiply imputed costs and utilities at each time point (raw values are provided in the electronic supplementary material [ESM]). Missing values were multiply imputed jointly for costs and utilities at each follow-up point using predictive mean matching and chained equations for 35 datasets given 35% loss to follow-up for complete cases. Baseline age and SF-12 Physical Component Summary score were identified

as predictors of missingness [10] and used in the imputations. Costs were grouped in three main categories: primary care contacts (GP, practice nurse, phlebotomist, other primary care contacts); antidepressant medications (summing ANTLER medications and other antidepressant medications as listed earlier); and CSRI-captured resource use (psychotherapy, other community-based contacts, emergency care). Only the CSRI-captured cost category could have missing values, as data on primary care contacts and antidepressant medication prescriptions were considered to be complete for all participants for whom data were available and were not imputed for the eight participants for whom we were unable to obtain any data from their electronic medical records. These eight participants all returned CSRI questionnaires at baseline with zero reported resource use, and six returned similar questionnaires with zero resource use at follow-up timepoints, whereas two did not return the CSRI questionnaires at 26 and 52 weeks.

The mean per-participant differences in 12-month costs and QALYs by randomised arm were jointly estimated from the 35 imputed datasets via bootstrapped seemingly unrelated regression with 100 iterations to account for the correlation between costs and QALYs [31, 32], adjusting for baseline values and the minimisation variables of study centre (four categories), antidepressant medication (four categories), and severity of depressive symptoms at baseline (binary variable indicating whether CIS-R score was above or below the latest calculated median at baseline), with imputed datasets combined according to Rubin's rules [33].

In line with recommendations made elsewhere [34, 35], we took a probabilistic approach to aid decision making for resource allocation and calculated the probability that discontinuation of antidepressants was cost effective for a range of thresholds of cost per QALY gained compared with antidepressant maintenance.

The incremental cost-effectiveness ratio (ICER) for each analysis was calculated as the mean estimated difference in costs divided by the mean estimated difference in QALYs, except where one arm was dominant. The bootstrapped results were plotted on cost-effectiveness planes (CEPs), and the proportions of estimates that were above the cost-effectiveness threshold were plotted on corresponding cost-effectiveness acceptability curves (CEACs) for a range of thresholds [34, 35].

2.4 Secondary and Sensitivity Analyses

ICERs, CEACs and CEPs are reported for the following secondary analyses.

1. Disease-specific health and social care costs using the EQ-5D-5L responses and crosswalk tariff [17] for the calculation of utilities and QALYs, as this is the EQ-5D-5L value set currently preferred by NICE.
2. Disease-specific health and social care costs using the SF-12 responses and SF-6D tariff for the calculation of utilities and QALYs [19] as this generic preference-based health-related quality-of-life measure has also been used extensively in the mental health context and is also acceptable to NICE.
3. Wider cost perspective including out-of-pocket and productivity costs and using the EQ-5D-5L responses and VSE tariff for the calculation of utilities and QALYs.
4. Wider cost perspective including out-of-pocket and productivity costs and using the EQ-5D-5L responses and crosswalk tariff [17] for the calculation of utilities and QALYs.
5. Wider cost perspective including out-of-pocket and productivity costs and using the SF-12 responses and SF-6D tariff for the calculation of utilities and QALYs [19].

The three wider cost perspective analyses listed above were included as these costs are potentially of interest in this disease context although not strictly required for inclusion in analyses for NICE. Further sensitivity analyses were conducted based on the primary health economic analysis (disease-specific health and social care cost perspective in England, and utilities calculated from EQ-5D-5L using the VSE tariff) for complete cases only using no imputation and bootstrapped seemingly unrelated regression with 1000 iterations, and for imputing zero cost for missing CSRI information (35 imputations and 100 bootstraps), to investigate the impact of these different ways of dealing with the missing data.

A post-hoc sensitivity analysis included relapse as a covariate at each follow-up point and for total costs and QALYs to investigate the relationship between relapse and costs and utilities, as the study team identified that this might be a more important factor than the randomised arm itself and could potentially be driving the observed results. This involved creating variables for each follow-up time point (3, 6, 9 and 12 months), which indicated whether or not participants had relapsed, as defined by the primary clinical outcome, at any time up to that time point. Relapses were assessed using a modified CIS-R assessment we call the retrospective CIS-R (rCIS-R) that enquires about depressive symptoms over the previous 12 weeks. The rCIS-R is a fully structured assessment that was self-administered on a computer. It asks the initial mandatory questions from the original CIS-R but asks patients whether they experienced depressive symptoms over the past 12 weeks (rather than the past week as in the original CIS-R). If participants answered positively to the mandatory questions, the subsequent questions in each section asked about the worst week during the

past 12 weeks. The rCIS-R was completed at every in-person follow-up except 6 weeks. Participants were asked to identify the number of weeks since the previous assessment when these symptoms began in order to estimate date of onset of relapse [10]. Stata v14 was used to run the analyses [36].

3 Results

In total, 478 participants were randomised between 9 March 2017 and 1 March 2019: $n = 238$ to maintenance antidepressant and $n = 240$ to discontinuation (see Section S1 of the ESM for the CONSORT diagram).

For participants randomised to maintenance antidepressant, 29% (70/238) were male, 93% (221/238) white, and the mean (standard deviation, SD) age was 54 (13) years. In the discontinuation arm, 25% (59/240) were male, 97% (228/235) were white, and the mean (SD) age was 55 (12) years. In the maintenance (discontinuation) arm, 47% (47%) were taking citalopram, 32% (35%) fluoxetine, 17% (15%) sertraline and 4% (3%) mirtazapine. The median time between randomisation and taking the study medication was 9 days (interquartile range [IQR] 6–13) in the maintenance arm and 8 days (IQR 6–13) in the discontinuation arm. Further details on the characteristics of trial participants are reported in the clinical-effectiveness paper [10].

3.1 Costs

Descriptive statistics for resource use are reported in the ESM (Section S2 for primary care costs, Section S3 for CSRI costs, Section S4 for antidepressant medications, including ANTLER medications, and Section S5 for the total cost statistics). Overall imputed costs for CSRI and adjusted costs for each resource use type are reported in Table 2.

Table 2 Total costs for primary care and other disease-specific health-related service use, over 12 months, adjusted for the baseline and minimisation variables

Variable	Maintenance group			Discontinuation group			Discontinuation vs. maintenance			
	n	Mean	SD/SE[a]	n	Mean	SD/SE[a]	Adjusted difference	Lower 95% CI	Upper 95% CI	p value
Total primary care cost										
Baseline	233	78.49	6.74	237	77.86	6.68				
6 months	233	90.81	7.14	237	98.57	7.08				
12 months	233	90.29	6.40	237	88.43	6.35				
Total (unadjusted)	233	181.24	228.17	237	186.86	163.63				
Total (adjusted)	233	181.10	11.58	237	187.00	11.48	5.90	− 25.39	37.19	0.71
Total CSRI costs (imputed)										
Baseline	232	49.07	16.89	236	11.26	16.75				
6 months	232	16.38	6.88	236	11.80	6.55				
12 months	232	16.05	5.25	236	19.48	5.24				
Total (unadjusted)	232	34.17	12.25	236	29.58	5.95				
Total (adjusted)	232	32.44	9.64	236	31.28	9.45	− 1.16	− 27.91	25.59	0.93
Antidepressant medications										
Baseline	238	5.41	0.20	240	5.38	0.20				
6 months	238	6.52	0.14	240	3.28	0.14				
12 months	238	6.69	0.22	240	3.90	0.22				
Total (unadjusted)	238	34.17	12.25	240	29.58	5.95				
Total (adjusted)	238	13.21	0.32	240	7.18	0.32	− 6.04	− 6.97	− 5.11	< 0.001
Total costs, all categories (primary care, CSRI imputed, meds)										
Baseline	232	132.99	18.02	236	94.53	17.87				
6 months	232	112.58	10.26	236	114.52	10.00				
12 months	232	112.14	8.76	236	113.31	8.70				
Total (unadjusted)	232	228.76	19.55	236	223.86	12.19				
Total (adjusted)	232	224.72	16.01	236	227.83	15.81	3.11	− 41.28	47.50	0.89

Costs are presented in £, year 2018/2019 values

CI confidence interval, *CSRI* Client Service Receipt Inventory, *SD* standard deviation, *SE* standard error

[a]SD when no multiple imputation was performed, SE when multiple imputation was performed

こ

Mean antidepressant medication costs per participant over 12 months were lower in the discontinuation arm than in the maintenance arm (mean per-participant difference of discontinuation minus maintenance of − £6.04; 95% confidence interval [CI] − 6.97 to − 5.11). There was a statistically significant difference between arms in GP consultation costs, where those in the discontinuation arm were a mean of £16 (95% CI 0.7–33) higher per participant than those in the maintenance arm over the course of 12 months, which equates to approximately half a GP visit. The discontinuation arm also had increased costs of psychological therapies over 12 months, with a mean additional cost per participant of £17 (95% CI 1.1–33), which equates to approximately 15–20 min of a therapist's time.

The numbers of participants reporting any psychotherapy use in the 18 months covered by data collection were as follows. In the 6-month period preceding baseline, 24/237 (10.1%) of people in the maintenance arm and 18/239 (7.5%) in the discontinuation arm reported some use of psychotherapy. In the first 6 months of the study follow-up period, the corresponding figures were 10/211 (4.7%) in the maintenance arm and 16/193 (8.3%) in the discontinuation arm. In the last 6 months of the study follow-up period, the figures were 15/210 (7.1%) in the maintenance arm and 30/181 (16.6%) in the discontinuation arm.

The mean (standard error, SE) total adjusted imputed health and social care costs were £228 (16) per participant in the discontinuation arm and £225 (16) per participant in the maintenance arm, with a mean adjusted bootstrapped difference over 12 months of £3 (95% CI − 41 to 48).

3.2 Utility Scores and Quality-Adjusted Life-Years

Mean adjusted imputed utility scores calculated from the EQ-5D-5L and VSE at each time point and 12-month QALYs are reported in Table 3, along with the same information for the EQ-5D-5L with crosswalk algorithm, and the SF-12 (SF-6D algorithm). The discontinuation arm had

Table 3 Descriptive statistics for utility scores at each timepoint and 12-month QALYs adjusting for baseline and minimisation variables

Variable	Maintenance group (0)			Discontinuation group (1)			Discontinuation vs. maintenance			
	n	Mean	SE	n	Mean	SE	Difference[a]	Lower 95% CI	Upper 95% CI	p value
EQ-5D-5L VSE										
Baseline	232	0.868	0.008	236	0.886	0.008	0.018	− 0.005	0.041	0.124
3 months	232	0.878	0.008	236	0.846	0.008	− 0.032	− 0.053	− 0.011	0.003
6 months	232	0.880	0.008	236	0.871	0.008	− 0.008	− 0.031	0.014	0.465
9 months	232	0.881	0.007	236	0.884	0.007	0.003	− 0.017	0.024	0.755
12 months	232	0.887	0.008	236	0.872	0.008	− 0.015	− 0.037	0.007	0.168
QALYs (unadjusted)[b]	232	0.876	0.119	236	0.869	0.115	− 0.019	− 0.035	− 0.003	0.020
QALYs adjusted[a]	232	0.880	0.005	236	0.869	0.005	− 0.011	− 0.026	0.003	0.127
EQ-5D-5L crosswalk										
Baseline	232	0.805	0.010	236	0.825	0.010	0.020	− 0.009	0.049	0.171
3 months	232	0.815	0.010	236	0.778	0.010	− 0.037	− 0.064	− 0.010	0.007
6 months	232	0.820	0.010	236	0.808	0.010	− 0.012	− 0.040	0.016	0.391
9 months	232	0.818	0.009	236	0.826	0.009	0.008	− 0.017	0.033	0.537
12 months	232	0.824	0.010	236	0.811	0.010	− 0.012	− 0.039	0.015	0.370
QALYs (unadjusted)[b]	232	0.814	0.151	236	0.806	0.139	− 0.022	− 0.042	− 0.002	0.028
QALYs (adjusted)[a]	232	0.818	0.006	236	0.806	0.006	− 0.012	− 0.030	0.006	0.187
SF-12/SF-6D										
Baseline	232	0.756	0.007	236	0.774	0.007	0.018	− 0.002	0.038	0.081
3 months	232	0.751	0.008	236	0.708	0.008	− 0.043	− 0.065	− 0.021	< 0.001
6 months	232	0.755	0.008	236	0.736	0.009	− 0.019	− 0.044	0.005	0.116
9 months	232	0.767	0.008	236	0.757	0.009	− 0.010	− 0.034	0.013	0.384
12 months	232	0.770	0.009	236	0.747	0.009	− 0.023	− 0.048	0.001	0.064
QALYs (unadjusted)[b]	232	0.759	0.103	236	0.743	0.100	− 0.027	− 0.043	− 0.011	0.001
QALYs (adjusted)[a]	232	0.760	0.005	236	0.739	0.006	− 0.021	− 0.036	− 0.006	0.006

CI confidence interval, *QALYs* quality-adjusted life-years, *SE* standard error, *SF-12* 12-Item Short Form survey, *VSE* value set for England

[a]Differences and adjusted means adjusted for baseline values (except at baseline) and minimisation variables

[b]Difference is unadjusted

a significantly lower utility at 3 months (difference using EQ-5D-5L VSE values − 0.032; 95% CI − 0.053 to − 0.011) and non-significantly fewer QALYs over the 12-month period (difference using EQ-5D-5L VSE values − 0.011; 95% CI − 0.026 to 0.003). There were no significant differences between arms in utilities at the other time points. The SF-6D QALYs showed a significant reduction in QALYs in the discontinuation group compared with the maintenance group (difference − 0.021; 95% CI − 0.036 to − 0.006). The trend in utilities can also be seen in Fig. 1, which shows the mean unadjusted complete-case utility scores at each time point for the three different methods of calculating utilities.

3.3 Cost-Utility Analysis

The overall result of a CUA is summarised as the ICER, which is the mean incremental cost per QALY gained of discontinuing antidepressant medication compared with maintenance antidepressant medication. In the primary analysis, with utilities and CSRI costs jointly imputed using multiple imputation with chained equations [37], then bootstrapping seemingly unrelated regression performed, discontinuation was dominated by maintenance in that it cost more (£2.71; 95% CI − 36.10 to 37.07) and resulted in fewer QALYs (− 0.010; 95% CI − 0.024 to 0.004) (see Table 4); however, the 95% CI crossed zero in both cases, so the differences were not statistically significant. These values differed slightly from those stated earlier because of the use of the seemingly unrelated regression, which accounts for the correlation between costs and utility scores so is a better estimate, but the conclusions reached are the same. The overall result is that the bootstrapped differences in costs and QALYs lie predominantly in the northwest quadrant of the CEP (see Fig. 2), suggesting that the discontinuation arm

is dominated, i.e. it incurs higher costs and provides fewer QALYs than maintenance, on average.

The information from the CEP was translated onto the CEAC (see Fig. 3), which shows the likelihood of discontinuation being cost effective at a range of values of the cost-effectiveness threshold. Figure 3 shows values up to £100,000/QALY. At the standard QALY threshold values of £20,000 and 30,000 per QALY gained, there was a 12.9 and 12.4% probability that discontinuation was cost effective compared with maintenance, respectively. The CEAC curve lies below 50% for all thresholds ≥ 0, in agreement with the conclusion that discontinuation is dominated by maintenance.

3.4 Secondary and Sensitivity Analyses

The results remained the same for all secondary analyses, including when productivity and out-of-pocket costs were included, i.e. discontinuation consistently resulted in higher costs and fewer QALYs than maintenance, on average. When considering wider societal costs, there was a £41 (95% CI − 222 to 303) adjusted difference in productivity loss costs for discontinuation compared with maintenance, with a total cost difference of £22 (95% CI − 179 to 219) when this, along with other private and out-of-pocket costs across the different costing categories, was added to the total health and social care costs and estimated jointly using bootstrapped seemingly unrelated regression.

When the SF-6D algorithm was used to calculate utilities from SF-12 responses, there were significantly fewer QALYs in the discontinuation arm (see Table 4).

For wider societal costs, we did not find a significant impact on productivity, partly because of the high variability in the reported numbers of days off. Tables showing the

Fig. 1 Unadjusted complete-case mean utility scores at each timepoint, by arm, for each of the three methods: EQ-5D-5L VSE, EQ-5D-5L crosswalk and SF-12/SF-6D. *SF-12* 12-Item Short-Form survey, *VSE* value set for England
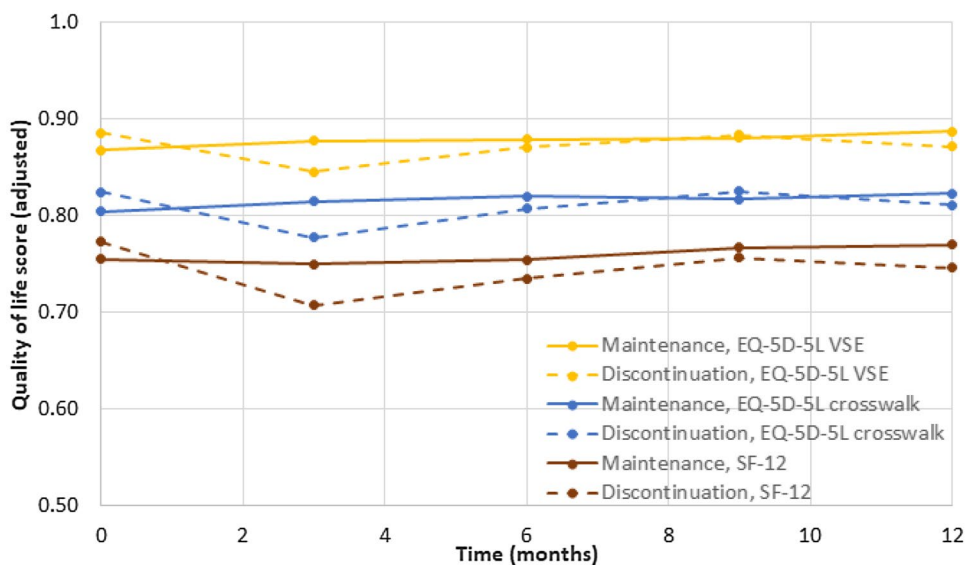
Table 4 Mean incremental costs[a], QALYs, ICERs and probabilities of the discontinuation arm being cost effective at the standard £20,000 and 30,000 per QALY gained thresholds commonly used by NICE

	Difference in costs	Lower 95% CI	Upper 95% CI	Difference in QALYs	Lower 95% CI	Upper 95% CI	ICER	Probability of discontinuation being cost effective at these thresholds (%)	
								£20,000/QALY	£30,000/QALY
Using MICE for missing utilities and CSRI costs; disease-specific health and social care costs									
EQ-5D-5L VSE	2.71	− 36.10	37.07	− 0.010	− 0.024	0.004	Maintenance dominates discontinuation	12.9	12.4
EQ-5D-5L crosswalk	4.19	− 31.42	36.40	− 0.012	− 0.028	0.004	Maintenance dominates discontinuation	11.2	11.0
SF-12/SF-6D	2.16	− 36.58	36.41	− 0.020	− 0.033	− 0.008	Maintenance dominates discontinuation	0.1	0.1
Using complete cases only; disease-specific health and social care costs									
EQ-5D-5L VSE	26.29	− 3.32	57.29	− 0.019	− 0.032	− 0.005	Maintenance dominates discontinuation	0.6	0.6
EQ-5D-5L crosswalk	26.22	− 3.61	57.29	− 0.022	− 0.038	− 0.006	Maintenance dominates discontinuation	1.4	1.3
SF-12/SF-6D	23.93	− 6.66	56.67	− 0.026	− 0.039	− 0.013	Maintenance dominates discontinuation	0.0	0.0
Using MICE for missing utilities and CSRI costs; wider costs including productivity losses and private costs									
EQ-5D-5L VSE	22.30	− 179.25	218.76	− 0.01	− 0.02	0.00	Maintenance dominates discontinuation	12.3	11.6
EQ-5D-5L crosswalk	21.46	− 168.86	210.15	− 0.01	− 0.03	0.00	Maintenance dominates discontinuation	13.6	11.8
SF-12/SF-6D	21.42	− 179.47	216.89	− 0.02	− 0.03	− 0.01	Maintenance dominates discontinuation	2.0	1.1
Using complete cases only; wider costs including productivity losses and private costs									
EQ-5D-5L VSE	105.99	− 128.37	331.22	− 0.02	− 0.03	− 0.01	Maintenance dominates discontinuation	2.3	1.2
EQ-5D-5L crosswalk	106.02	− 128.38	331.39	− 0.02	− 0.04	− 0.01	Maintenance dominates discontinuation	2.4	1.6
SF-12/SF-6D	101.48	− 142.45	341.27	− 0.03	− 0.04	− 0.01	Maintenance dominates discontinuation	0.3	0.1

Results shown for the base-case analysis and all secondary analyses in terms of using different quality-of-life information and value sets and using different cost perspectives

CI confidence interval, *CSRI* Client Service Receipt Inventory, *ICER* incremental cost-effectiveness ratio, *MICE* multiple imputation using chained equations, *NICE* National Institute for Health and Care Excellence, *QALYs* quality-adjusted life-years, *SF-12* 12-Item Short Form survey, *VSE* value set for England

[a]Costs are presented in £, year 2018/2019 values
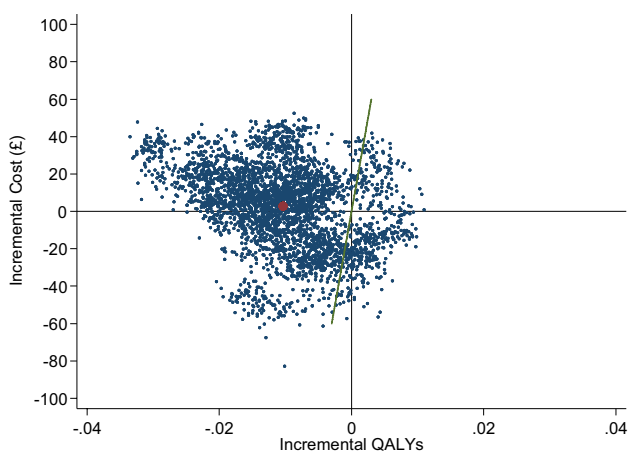
Fig. 2 Cost-effectiveness plane using multiple imputation using chained equations for utilities and Client Service Receipt Inventory costs, showing 100 bootstrapped results for each of 35 imputed datasets from using seemingly unrelated regression for disease-specific health-related costs and QALYs (EQ-5D-5L VSE). *Red point* is the mean costs vs. mean QALYs. *Straight line* is the £20,000/QALY gained threshold. Costs in £, year 2018–19 values. *QALY* quality-adjusted life-year, *VSE* value set for England
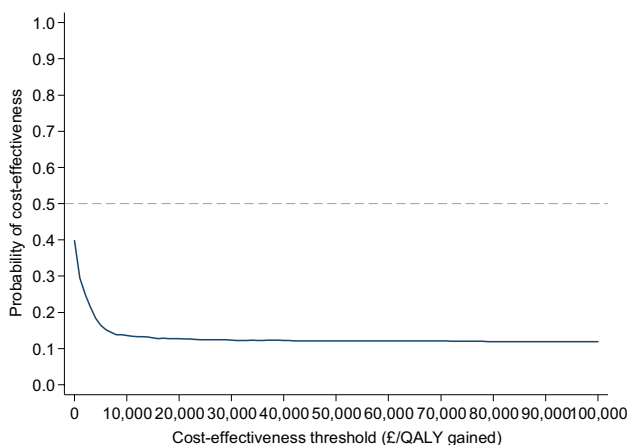
Fig. 3 Cost-effectiveness acceptability curve generated from the cost-effectiveness plane in Fig. 2 showing the probability of the discontinuation arm being cost effective compared with the maintenance arm, at a range of values for the cost-effectiveness threshold from £0 to 100,000/QALY. Costs in £, year 2018–19 values. *QALY* quality-adjusted life-year

numbers of days off work are given in Section S8 of the ESM.

When relapse was included in the adjusted bootstrapped regression analyses, the difference in utilities at 3 months was significant for both relapse (− 0.053; 95% CI − 0.079 to − 0.028) and randomised arm (− 0.026; 95% CI − 0.047 to − 0.005) (values are given for the EQ-5D-5L VSE here, but the differences were also significant for the other QOL

methods—see Section S7 of the ESM). With the difference in QALYs over 12 months, there was a significant difference for relapse (− 0.046; 95% CI − 0.060 to − 0.032) and not for the randomised arm (− 0.002; 95% CI − 0.017 to 0.012). The main clinical analyses [10] showed that the time to relapse was significantly longer in the maintenance arm than in the discontinuation arm. Tables showing further details of the utility and QALY results when considering the relapse status can be found in Section S7 of the ESM.

Relapse also had a significant impact on the cost of GP appointments (relapse cost an additional £34 [95% CI 17–51], adjusted total over 12-month period), and the randomised arm was no longer significant, which also followed through into a significant difference in total primary care costs between those who relapsed (more expensive by £50 [95% CI 18–83] over the year) and those who did not. The difference in antidepressant medication cost was not explained by relapse status, only by randomised arm. The total cost (primary care contacts, medications and imputed CSRI-collected costs) was significantly different according to relapse status, with those who relapsed costing £70 (95% CI 25–115) more overall over the year than those who did not. Tables showing these values are in Sections S3–S6 of the ESM. CEPs and CEACs for the secondary analyses are presented in Section S9 of the ESM.

4 Discussion

4.1 Main Findings

The CUA described in this paper suggests that, over the 12-month period of the ANTLER study, there was a low probability that tapering of the antidepressant medications citalopram, sertraline, fluoxetine or mirtazapine was cost effective compared with continued maintenance of antidepressant treatment in this population. Participants randomised to the discontinuation arm had significantly lower utility scores at 3 months, most likely driven by an increased probability of depression relapse compared with antidepressant maintenance, although there was no significant difference for 12-month QALYs for the EQ-5D-5L VSE or for the EQ-5D-5L crosswalk. However, the SF-6D as generated from the SF-12 did show a significant difference for 12-month QALYs. Participants randomised to discontinuation also had greater GP consultation and psychotherapy costs. This similarly appeared to be driven by a shorter time to relapse in the discontinuation arm for GP costs, potentially as participants arranged to see their GP following relapse to review their medication, with 53% (95% CI 44–62) of those who relapsed in the discontinuation arm

having returned to a known antidepressant before the end of the trial [10].

4.2 Strengths and Weaknesses

Although a strength of the analysis is our relatively complete dataset for medications and primary care costs, as these were obtained from primary care electronic medical records, we had a 20% loss to follow-up for self-reported questionnaires, which has implications for the calculation of QALYs in particular. While there was no evidence that follow-up data were missing not at random, there is generally a risk of bias when data are missing. A further limitation of the analysis was that the effectiveness measures used in our analysis were generic health-related quality-of-life measures. Current evidence suggests that the EQ-5D-5L valued using the VSE is responsive to symptomatic changes in depression, but it remains possible that information regarding other factors that are important to participants were not captured, particularly factors related to recovery from mental illness that are included in the new disease-specific ReQoL (Recovering Quality of Life) measure [24]. Notably, the SF-12/SF-6D analyses did show a greater sensitivity in this population, as has been reported in the literature [38], potentially because it includes more questions around emotions. None of these measures capture patient preferences for treatment modality, and they may all be limited in how well they capture the side effects of antidepressant medications. Discrete choice experiments are a potential methodology for capturing patient preferences for depression treatment outcomes, with evidence in this area currently limited [39].

In the CSRI, patients were only asked about resource use related to their mental health, so any use of services outside this categorisation would not have been captured. Overall, there was limited use of secondary care services, particularly inpatient stays, with only a single inpatient stay reported.

Other than a comparison of patients who did and did not experience relapse during the study, we did not identify any suitable subgroup analyses as part of this study. Future research should consider how specific patient characteristics might interact with antidepressant discontinuation and the probability of relapse.

4.3 Implications

Given the differences in QALYs and costs between discontinuation and maintenance, it seems unlikely that discontinuation could potentially become cost effective over the longer term. At a £20,000–30,000 threshold for a QALY gain (i.e. the preferred NICE threshold), there would need to be a cost saving of £200–600 per person for discontinuation over the life-time horizon after the first 12 months to balance the utility lost during the first 3 months after discontinuing the medication, and it is unclear where in the longer-term patient pathway this could potentially occur. Trials and decision analytic models of depression rarely go beyond 12 months [4–6], so any longer-term modelling to try and answer this question would require untested assumptions. This highlights the need for further research into the longer-term implications of discontinuing antidepressant use compared with continuing with long-term maintenance for the prevention of relapse, in terms of both longer-term costs and longer-term benefits or harms to patients.

This analysis considers the discontinuation of an existing treatment, rather than the initiation of a new treatment. The more common context of an economic evaluation alongside a clinical trial is that of adding a new technology and then evaluating whether the potential increase in costs to the NHS associated with using the new technology is justified in terms of increased QALYs for patients or cost savings seen elsewhere. However, in this analysis, we were interested in whether discontinuing a treatment had an adverse impact on patients and whether any adverse impact could be offset by sufficiently large cost savings.

In economic evaluations, we also tend to compare a new technology with the current gold standard of care. In this situation, there is some question regarding which regimen is standard of care and which is the novel intervention. Although NICE guidance suggests that patients should stay on antidepressants for at least 6 months after remission, there is limited guidance on when antidepressants should be discontinued, and hence which regimen should be the standard of care in this comparison. For the patients recruited to ANTLER, standard of care was long-term maintenance of antidepressants.

The analysis showed no evidence of cost savings for the NHS on discontinuing treatment, and a potential disadvantage to patients' health-related quality of life in the short term. However, this difference in health-related quality of life between the two treatment pathways disappeared by 12 months, meaning that any detriment due to relapse or to discontinuing the original medication, while statistically significant, was on average short lived, due in part to some patients resuming taking antidepressants [10].

It was not possible to include information on whether the potential disadvantage suggested here is meaningful for patients in the context of other concerns that were not captured by our analysis, such as feelings of stability derived from maintenance medication, or feelings of liberation derived from having tapered to a zero dose and become 'antidepressant free' and therefore also free of medication side effects such as weight gain, sleep disturbance and sexual dysfunction and potential longer-term harms arising from taking these medications. Surveys of long-term antidepressant users found that between two-thirds and three-quarters of patients reported some adverse effects from antidepressant use, with over two-thirds also

reporting that antidepressants helped them get by or cope. Information on the long-term risks of taking antidepressant medication and increased clinical support for discontinuation are both important to patients on maintenance antidepressants [40]. Discontinuation is also likely to occur alongside non-pharmacological treatments. However, evidence for the most clinically and cost-effective way to do this remains limited [41], with non-pharmacological treatments such as talking therapies requiring more resources and potentially being harder for patients to access. It would be useful if future work could investigate predictive characteristics regarding which patients might be more likely to relapse on discontinuing long-term antidepressant medication.

In summary, this analysis provides information to inform guidance on joint decision making between the patient and their clinician regarding their continued antidepressant prescription and would not support a change in NICE guidance to advise that all patients on maintenance antidepressants should follow a single preferred pathway.

5 Conclusions

This study reports important new evidence regarding the health economic considerations of decisions regarding discontinuing long-term maintenance antidepressants in currently well patients. Based on the results of this study, discontinuation of medications would not be recommended nationally on the grounds of cost effectiveness. Despite this, some individuals may choose to taper and stop antidepressants to see whether they can manage without antidepressants, as they may have other important and influential considerations that were not captured as part of this study.

Declarations

Conflict of interest TK received Grants from NIHR during the conduct of the study. RH, NW and MM received Grants from NIHR HTA during the conduct of the study. GL received grants from UCL during the conduct of the study and personal fees from Fortitude Law, outside the submitted work. NF has received personal fees from ALK, Allergan, Aimmune, AstraZeneca, MSD, Ipsen, Sanofi Aventis, Novo Nordisk, Grifols, Vertex and Abbott Singapore and grants from the European Association of Cardiothoracic Surgery, outside the submitted work. CSC, LD, SG, DK, MK, PL, DM, IN and LM have no conflicts of interest that are directly relevant to the content of this article.

Availability of data and material De-identified individual participant data and analytical code will be available from the authors to bona fide investigators for research projects that have been approved by independent review committees, after signing a data sharing agreement with the study sponsor (University College London, London, UK).

Ethics approval Ethical approval was obtained from the National Research Ethics Service committee, East of England, Cambridge South (ref.: 16/EE/0032). Clinical trial authorisation was given by the MHRA. The trial sponsor was University College London. The trial was registered: EudraCT number 2015-004210-26; protocol number 14/0647 (version 7.0); Controlled Trials ISRCTN Registry, ISRCTN15969819.

Consent All study participants provided written informed consent before participating in the trial.

Author contributions CSC was lead author, and wrote the first draft of the article. CSC and RH planned and conducted the economic analyses. All authors were involved in formulating the overall research question, including economic aspects, and designing and conducting the study. PL was patient and public involvement collaborator. CSC, RH, LM, NF, GL and LD designed the economic data collection instruments. LD managed the data collection. LD and GL managed the trial. CSC, LM and LD were involved in data cleaning. CSC, RH and LM analysed the data. SG, TK, NW and DK were involved in participant recruitment. GL, IN, MK, DK, NW, SG, NF, RH, TK and MM were co-applicants on the funding application. All authors contributed to and approved the final manuscript.

References

1. Viola S, Moncrieff J. Claims for sickness and disability benefits owing to mental disorders in the UK: trends from 1995 to 2014. Br J Psychiatry Open. 2016;2(1):18–24.

2. McCrone P, Dhanasiri S, Patel A, Knapp M, Lawton-Smith S. Paying the Price: the cost of mental health care in England to 2026. London: King's Fund; 2008.

3. McCrea RL, Sammon CJ, Nazareth I, Petersen I. Initiation and duration of selective serotonin reuptake inhibitor prescribing over time: UK cohort study. Br J Psychiatry. 2016;209(5):421–6.

4. NICE. Depression in adults: recognition and management; clinical guideline CG90. 2009. [Online]. Available: https://www.nice.org.uk/guidance/cg90.

5. Ramsberg J, Asseburg C, Henriksson M. Effectiveness and cost-effectiveness of antidepressants in primary care: a multiple treatment comparison meta-analysis and cost-effectiveness model. PLoS ONE. 2012;7(8):e42003.

6. Annemans L, Brignone M, Druais S, De Pauw A, Gauthier A, Demyttenaere K. Cost-effectiveness analysis of pharmaceutical treatment options in the first-line management of major depressive disorder in Belgium. Pharmacoeconomics. 2014;32:479–93.

7. Thom H, Jackson C, Welton N, Sharples L. Using parameter constraints to choose state structures in cost-effectiveness modelling. Pharmacoeconomics. 2017;35:951–62.

8. Marsden J, White M, Annand F, Burkinshaw P, Carville S, Eastwood B, Kelleher M, Knight J, O'Connor R, Tran A, Willey P, Greaves F, Taylor S. Medicines associated with dependence or withdrawal: a mixed-methods public health review and national database study in England. Lancet Psychiatry. 2019;6(11):935–50.

9. Duffy L, Bacon F, Clarke CS, Donkor Y, Freemantle N, Gilbody S, Hunter RM, Kendrick T, Kessler D, King M, Lanham P, Lewis G, Mangin D, Marston L, Moore M, Nazareth I, Wiles N, Lewis G. A randomised controlled trial assessing the use of citalopram, sertraline, fluoxetine and mirtazapine in preventing relapse in primary care patients who are taking long-term maintenance antidepressants (ANTLER: ANTidepressants to prevent reLapse in dEpRes. Trials. 2019;20:319.

10. Lewis G, Marston L, Duffy L, Freemantle N, Gilbody S, Hunter R, Kendrick T, Kessler D, Mangin D, King M, Lanham P, Moore M, Nazareth I, Wiles N, Bacon F, Bird M, Brabyn S, Burns A, Clarke CS, Hunt A, Pervin J, Lewis G. Maintenance or discontinuation of antidepressants in primary care. N Engl J Med. 2021;385(14):1257–67.

11. Taves DR. Minimization: a new method of assigning subjects to treatment and control groups. Clin Pharmacol Ther. 1974;15:443–53.

12. Joint Formulary Committee. British National Formulary (online). [Online]. Available: http://www.medicinescomplete.com. Accessed Feb 2020.

13. Beecham J, Knapp M. Costing psychiatric interventions. In: Thornicroft G, editor. Measuring mental health needs. London: Gaskell; 2001. p. 200–24.

14. Curtis L, Burns A. Unit costs of health and social care 2019. Canterbury: Personal Social Services Research Unit, University of Kent; 2019.

15. NHS Improvement. NHS reference costs. [Online]. Available: https://improvement.nhs.uk/resources/national-cost-collection/.

16. Office for National Statistics. ONS occupation codes (PROV—Occupation SOC10 (4) Table 14.1a Weekly pay—Gross 2019.xls). 29 October 2019. [Online]. Available: https://www.ons.gov.uk/employmentandlabourmarket/peopleinwork/earningsandworkinghours/datasets/occupation4digitsoc2010ashetable14. Accessed June 2020.

17. Dolan P. Modelling valuations for EuroQol health states. Med Care. 1997;35:1095–108.

18. Ware J Jr, Kosinski M, Keller SD. A 12-Item Short-Form Health Survey: construction of scales and preliminary tests of reliability and validity. Med Care. 1996;34(3):220–33.

19. Brazier JE, Roberts J. The estimation of a preference-based measure of health from the SF-12. Med Care. 2004;42(9):851–9.

20. Devlin NJ, Krabbe PF. The development of new research methods for the valuation of EQ-5D-5L. Eur J Health Econ. 2013;14(Suppl. 1):S1–3.

21. Devlin NJ, Shah KK, Feng Y, Mulhern B, van Hout B. Valuing health-related quality of life: an EQ-5D-5L value set for England. Health Econ. 2018;27:7–22.

22. van Hout B, Janssen MF, Feng Y-S, Kohlmann T, Busschbach J, Golicki D, Lloyd A, Scalone L, Kind P, Pickard AS. Interim scoring for the EQ-5D-5L: mapping the EQ-5D-5L to EQ-5D-3L value sets. Value Health. 2012;15(5):708–15.

23. NICE. Guide to the methods of technology appraisal. 2013. [Online]. Available: https://www.nice.org.uk/process/pmg9.

24. Franklin M, Enrique A, Palacios J, Richard D. Psychometric assessment of EQ-5D-5L and ReQoL measures in patients with anxiety and depression: construct validity and responsiveness. Qual Life Res. 2021;30:2633–47.

25. Marston L, Freemantle N, Hunter RM, Clarke CS, Duffy L, Lewis G. ANTLER statistical analysis plan, version 3, 16 January 2020. UCL Discovery database.

26. Isaacs A, Critchley J, See Tai S, Buckingham K, Westley D, Harridge SD, Smith C, Gottlieb JM. Exercise Evaluation Randomised Trial (EXERT): a randomised trial comparing GP referral for leisure centre-based exercise, community-based walking and advice only. Health Technol Assess. 2007;11(10). https://doi.org/10.3310/hta11100.

27. Pope I, Burn H, Ismail SA, Harris T, McCoy D. A qualitative study exploring the factors influencing admission to hospital from the emergency department. BMJ Open. 2017;7:e011543.

28. Rosie Lovett and Sophie Cooper, NICE Science policy and research programme. NICE to support new valuation study for England for EQ-5D-5L questionnaire; study will measure public's preference for different health states. 2019. [Online]. Available: https://www.nice.org.uk/news/blog/nice-to-support-new-valuation-study-for-england-for-eq-5d-5l-questionnaire.

29. Hunter RM, Baio G, Butt T, Morris S, Round J, Freemantle N. An educational review of the statistical issues in analysing utility data for cost-utility analysis. Pharmacoeconomics. 2015;33(4):355–66.

30. Kearns B. NICE DSU report: the relevance of future, unrelated health costs in economic evaluation in NICE appraisals. Decision Support Unit, School of Health and Related Research (ScHARR), University of Sheffield; 2020.

31. Zellner A. An efficient method of estimating seemingly unrelated regression equations and tests for aggregation bias. J Am Stat Assoc. 1962;57(298):348–68.

32. Leurent B, Gomes M, Faria R, Morris S, Grieve R, Carpenter JR. Sensitivity analysis for not-at-random missing data in trial-based cost-effectiveness analysis: a tutorial. Pharmacoeconomics. 2018;36:889–901.

33. Rubin DB. Multiple imputation for nonresponse in surveys. New York: Wiley; 1987.

34. Fenwick E, Claxton K, Sculpher M. Representing uncertainty: the role of cost-effectiveness acceptability curves. Health Econ. 2001;10:779–87.

35. Briggs AH, Wonderling DE, Mooney CZ. Pulling cost-effectiveness analysis up by its bootstraps: a non-parametric approach to confidence interval estimation. Health Econ. 1997;6:327–40.

36. StataCorp. Stata Statistical Software: release 14. College Station: StataCorp LP; 2015.

37. Royston P, White I. Multiple imputation by chained equations (MICE): implementation in stata. J Stat Softw. 2011;45(4).

38. McCrone P, Patel A, Knapp M, Schene A, Koeter M, Amaddeo F. A comparison of SF-6D and EQ-5D utility scores in a study of patients with schizophrenia. J Ment Health Policy Econ. 2009;12(1):27–31.

39. Morey E, Thacher J, Craighead WE. Patient preferences for depression treatment programs and willingness to pay for treatment. J Ment Health Policy Econ. 2007;10(2):73–85.

40. Cartwright C, Gibson K, Read J, Cowan O, Dehar T. Long-term antidepressant use: patient perspectives of benefits and adverse effects. Patient Prefer Adher. 2016;10:1401–7.

41. Kuyken W, Hayes R, Barrett B, Byng R, Dalgleish T, Kessler D, Lewis G, Watkins E, Brejcha C, Cardy J, Causley A, Cowderoy S, Evans A, Gradinger F, Kaur S, Lanham P, Morant N, Richards J, Shah P, Sutton H, Vicary R, Weaver A, Wilks J, Williams M, Taylor RS, Byford S. Effectiveness and cost-effectiveness of mindfulness-based cognitive therapy compared with maintenance antidepressant treatment in the prevention of depressive relapse or recurrence (PREVENT): a randomised controlled trial. Lancet. 2015;386:63–73.

Authors and Affiliations

Caroline S. Clarke[1] [ORCID] · **Larisa Duffy[2]** · **Glyn Lewis[2]** · **Nick Freemantle[3]** · **Simon Gilbody[4]** · **Tony Kendrick[5]** · **David Kessler[6]** · **Michael King[2]** · **Paul Lanham[7]** · **Derelie Mangin[8]** · **Michael Moore[5]** · **Irwin Nazareth[1]** · **Nicola Wiles[9]** · **Louise Marston[1]** · **Rachael Maree Hunter[1]**

1 Research Department of Primary Care and Population Health, University College London, London, UK

2 Division of Psychiatry, University College London, London, UK

3 Institute of Clinical Trials and Methodology, University College London, London, UK

4 Department of Health Sciences and Hull York Medical School, University of York, York, UK

5 Primary Care Population Sciences and Medical Education, Faculty of Medicine, University of Southampton, Southampton, UK

6 Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK

7 Patient and Public Involvement Collaborator, London, UK

8 Department of Family Medicine, McMaster University, Hamilton, Canada

9 Centre for Academic Mental Health, Population Health Sciences, Bristol Medical School, University of Bristol, Bristol, UK