HEALTH ECONOMICS ANALYSIS PLAN

Version: 2.0

Contents

SECTION 1: ADMINISTRATIVE INFORMATION

Title: Chronic Headache and Self-management Study (CHESS)

ISRCTN number: *79708100*

HEAP Version: Version 2.0 (Date: 27.Nov.2019

Protocol Version: Version 3.7 (Date: 19.09.2019

HEAP revisions: None

Roles and responsibility:

Signatures of:

(signed off on Q Pulse at dates below)

	Name	Date	Signature
Author of HEAP	Felix Achana	27/11/2019	
Co-applicant	Stavros Petrou	27/11/2019	
Co-applicant	Hema Mistry	22/01/2020	
Chief Investigator	Martin Underwood	24/01/2020	

SECTION 2: INTRODUCTION

Objective

The Chronic Headache Education and Self-management Study (CHESS) is a multicomponent programme of interlocking studies funded by an NIHR programme grant to develop an education and self-management support intervention for people living with chronic headache (here in referred to as the CHESS intervention) and assess its clinical and cost-effectiveness in a randomised controlled trial. This analysis plan relates to the economic evaluation of the CHESS intervention using data from the two-arm multi-centre randomised controlled trial component of the CHESS programme. The within-trial economic evaluation will aim to estimate the cost-effectiveness of the CHESS intervention compared with best supportive care over the 12-month trial period of follow-up. The purpose of the health economics analysis plan is to outline an explicit framework of methods that will be used to analyse the health economic data in a robust manner. The document has been written based on information contained in the trial protocol version 3.7 dated on 19.Sep.2019.

Background rationale

Chronic headaches present a major problem both for the individual and society. Previous studies on supportive self-management interventions in this population have largely been small studies with short term follow-up, they often did not report clinically relevant outcomes, or were conducted in different healthcare systems therefore difficult to translate into an NHS setting. These studies also did not necessarily focus on chronic headache but rather looked at headache with no frequency specified. Based on the results of our systematic review there may be potential for large gain through a combination of self-

management education and appropriate use of prophylaxis and management of medication overuse headache in a chronic headache population.

In order to develop the evidence base needed for self-management intervention for chronic headache there needs to be a carefully developed, piloted and evaluated intervention package which has been supported by good qualitative work on understanding outcomes of interest. There is therefore the need for a robust clinical and cost-effectiveness trial within an NHS setting.

Objectives

The objective is to answer the question: Amongst adults with chronic headache arising from migraine, chronic tension type headache or medication overuse headache, is the provision of a self-management support programme in addition to best usual NHS care clinically and cost effective?

SECTION THREE: METHODS

General principles for economic evaluation

The within-trial economic analysis will be conducted under the intention to treat (ITT) principle. This requires that study participants are analysed according to their treatment assignment regardless of actual treatment received (1). The perspective of the base case analysis will be that of the UK National Health Service and Personal Social Services (NHS/PSS), the recommended perspective for technology appraisals in the National Institute for Health and Care Excellence (NICE) reference case (2). Secondary analyses will consider costs from a wider societal perspective (3). A 12-month time horizon will be adopted for the within-trial analysis to mirror the trial follow-up period and therefore costs and outcomes will not be discounted due to this shorter time horizon. However, we will develop a decision analytic model to extrapolate trial results beyond the trial follow-up and assess the long-term cost-effectiveness of the CHESS intervention. Costs and outcomes in the decision model will be discounted at 3.5% beyond the first year post randomisation in accordance

with the NICE reference case (2). Findings will be reported in accordance with the Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement for the reporting of health economic evaluations (4).

Resource use and costs

Health and social care resource use will be collected for each trial participant over the 12-month period of follow-up. As outlined in the study protocol, the CHESS intervention consists of i) a structured education and self-management sessions delivered to groups of 10-12 patients over two days, ii) one-to-one consultation with the group facilitator (usually a registered nurse) for each participant and iii) a follow-on telephone call within the first 8 weeks of participating in the group session. Each interventional group will be facilitated by a trained registered nurse (grade 5 and above) and one allied health professional. Resource use and costs associated with delivery of the intervention will be estimated based on: (i) a detailed record of each group activity including the number of patients attending each group, duration of sessions, number of staff facilitators and their respective grades and set-up costs such as administrative support, educational material/leaflets and the room/facilities where the group activities takes place, (ii) number and length of one-to-one consultations with the clinical nurse and (iii) number and length of telephone follow-up consultations and the clinical grade of the staff conducting the consultation. Participants in the control group will be provided with a relaxation CD, the unit cost of which will be calculated based on the procurement costs for use of the CD in the trial.

In addition to the resource use associated with delivery of the interventions, resource utilisation data covering the 4-month period prior to randomisation (to establish baseline estimates) and the 12-month post-randomisation period will be collected for each trial participant through two principal means: (i) the trial case report forms including relevant primary/community care service use and hospital inpatient admissions and outpatient attendances and (ii) the computerised electronic record systems of participating general practice (GP) surgeries. Primary care, hospital inpatient and outpatient resource utilisation will be extracted from these sources for each trial participant. Primary care utilisation will be extracted from the electronic general practice records, which include details of consultations i.e. the number and type of consultations for example with a GP, practice

nurse or other community based health and social care professional or service and prescriptions. Secondary care utilisation data to be extracted from the GP electronic health records will include details of hospital day case and inpatient admissions (referral method and type of admission, type of ward, length of stay and details of diagnosis and procedures undertaken) and details of outpatient attendances (for example, headache clinic/neurology clinics, physiotherapy clinics, accident and emergency, medical tests, scans and investigations). Economic questionnaires completed by study participants at 4, 8 and 12 month assessment will provide additional secondary sources of NHS and Personal social service utilisation (community health and social care encounters and utilisation of hospital-based services). Costs based on resource use extracted from the GP records will act as the primary source of cost data for the economic evaluation. Costs estimated from resource use collected through the patient reported questionnaires will act as secondary data sources and will only be used where no equivalent cost information is available from the GP records. Private healthcare utilisation (including over the counter medication use), out-of-pocket expenses and travel costs borne by participants and their relatives, time-off work due to illness, lost income and use of community social care services such as meals on wheels (although use of these would most likely be minimal for the CHESS trial population). These will be measured using the economic questionnaires completed by study participants at 4, 8 and 12 month assessment. Private healthcare costs will be categorised into costs borne by other sector of the economy, e.g. use of community social care services, and cost borne by individuals.

Current UK unit costs will be applied to each resource item to value total resource use in each arm of the trial. A per diem cost for each level of hospital care, delineated by level of intensity, will be calculated using national tariffs. The unit costs of community health and social services will largely be derived from latest Unit Costs of Health and Social Care 2018 report published by the Personal Social Services Research Unit (PSSRU)(5), supplemented by information obtained from published literature and online sources. The primary analysis will concentrate on direct intervention and broader healthcare/PSS costs, whilst wider impact (societal) costs will be included within one of the sensitivity analyses.

Outcomes

The primary outcome of the within-trial economic evaluation will be the quality-adjusted life year (QALY) as recommended in the NICE reference case (2). This will allow incremental cost-effectiveness ratios for CHESS intervention compared with best usual care to be generated in the form of incremental cost per QALY gained. The QALY is a measure that combines quantity and quality of life lived into a single metric, with one QALY notionally equating to one year of full health. QALY estimates are generated from combining length and health-related quality of life outcomes using area-under-the-curve approaches (6). This requires survival and health-related quality of life data from or on behalf of trial participants for the period covering the trial time horizon. Health-related quality of life collected for trial participants (see details below) will be converted into health-state utilities indexed at 0 and 1 where 0 represents death and 1 represents full health.

Participants will be asked to complete the EuroQoL EQ-5D-5L (7) and SF-12 (8) measures using postal questionnaire at baseline and during follow-up at the 4, 8 and 12 months post-randomisation assessment points. Responses to the EQ-5D and SF-12 will be converted into multi-attribute utility scores using established algorithms (9, 10) from which QALYs can be generated. The EQ-5D is a generic preference based 5-dimensional multi-attribute instrument for measuring health-related quality of life. Currently, there are two versions of the questionnaire: a 3-level version (EQ-5D-3L) first introduced in 1990 by the EuroQoL Group (11) and a newer 5-level version (EQ-5D-5L) introduced in 2009 (12). Patients in the CHESS trial will complete the 5L version of the questionnaire. The 5L responses can be converted into health utilities using a recently published value set for England (13). However, since publication of the EQ-5D-5L value set, NICE has released a position statement (14) advising against the use of the new tariff (13) until the outcome of ongoing research exploring the impact of adopting the EQ-5D-5L valuation set in the NICE reference case becomes available. The position statement further recommends that during this interim period, EQ-5D-5L responses should be mapped or cross-walked onto the EQ-5D-3L using the Hout et al. (15) algorithm and the health utilities then derived from EQ-5D-3L utility scores using the UK value set for the EQ-5D-3L (16). Therefore, we initially plan to use the utility values derived from cross-walking the EQ-5D-5L responses onto the EQ-5D-3L

using the Hout et al. method to generate QALYs for the base case analysis. Sensitivity analyses will s will be conducted using health utility values generated from the SF-12 using the algorithm of Brazier et al (17).

SECTION FOUR: Mapping sub-study

A separate sub-study will be conducted as part of the CHESS programme of research to develop methods for mapping or cross-walking two headache-specific questionnaires (the 6-item Headache Impact Test (HIT-6) and the Chronic Headache Quality of Life Questionnaire (CHQLQ v1) onto generic health related quality of life questionnaires (the EQ-5D-5L and the SF-12 v2). A cross-sectional sample (sample size: 400-500) of people living with chronic headaches will be recruited from among patients attending headache clinics within NHS hospital outpatient departments for the mapping study. The headache-specific questionnaires are more likely to be responsive to improvement or worsening in headache-related symptoms than generic health-related quality of life measures such as the EQ-5D-5L and SF-12. Utilities based on the EQ-5D-5L or the SF-6D (via SF-12) can then be derived from the mapping algorithms. We will use utilities generated from the HIT-6 and the CHQLQ via the mapping functions as an alternative source of health utility in the base-case analysis where data from the EQ-5D-5L and the SF-12 v2 are missing. We will also use them stand alone sensitivity analyses to explore the robustness of the cost-effectiveness results to different approaches to measuring health-related quality of life impact of intervention.

SECTION FIVE: DATA

Data quality and cleaning

All data relevant to the health economics analysis will be examined for data quality. Questionnaires will be checked for completeness on return to the trial office. Any questionable data will be queried with trial staff and inappropriate or unclear responses will be handled in accordance with pre-specified data entry guidance. Unresolved issues after referral to the data entry instructions will be discussed with the trial health economists and clarification sought from the clinical team if necessary. Agreed line of actions for addressing data quality issues will be documented in the data entry guidance documentation.

Missing data

Any missing items present after the data cleaning stage will be addressed within the health economic analysis strategy as missing data. Missing data is a common occurrence within trial-based economic evaluations and it is necessary to address it in a standardised principled manner. Within the health economic literature, trial-based economic evaluations have been subject to particular criticism for failing to use appropriate methods to address missing data (18). Descriptive analyses of missing data will be carried out (missing data patterns using graphical tools, association between missing data and baseline variables, association between missing data and outcomes). The results of the descriptive analysis will be discussed by the trial team to infer possible reasons for missing data and inform the assumption about the missing data mechanism. In line with best practice recommendations for analysis of within-trial economic data (19), multiple imputation by chain equations implemented through the MICE package (20) in statistical package R version 3.13 (21) will be used to handle missing data for each assessment point (baseline, 4-, 8- and 12-month follow-up). Multiple imputation (MI) generates a series of datasets with each dataset replacing missing values with sampled values. MI replaces each missing observation with a set of plausible imputed values, taken from the predictive distribution of the missing data given the observed data (22). Such methods can handle data assumed missing at random (MAR) and can be modified to handle data assumed missing not at random (MNAR) (23). Appropriateness of the MAR assumption will be assessed by comparing the characteristics of patients with and without missing data at each follow-up time point. Imputated data will be generated separately by treatment group as recommended by Faria et al (24) using the predictive mean matching method which has the advantage of preserving non-linear relationships and correlations between variables within the data. Estimates obtained will be pooled to generate mean and variance estimates of costs and QALYs using Rubin's rule in order to capture within and between variances for imputed samples. We will fit models under a missing not at random (NMAR) assumption by systematically varying values of imputed costs and utilities from 0 to ±100% within the imputation models to assess the robustness of our base-case results to the missing at random assumption.

SECTION SIX: ANALYSIS

Summary of resource use and costs

Patient-level costs will be generated for each resource variable by multiplying the quantity reported with the respective unit cost, weighted by length of stay or duration of contact where appropriate. Summary statistics (means, standard errors and completion rates) will be generated by treatment allocation and assessment point. Between treatment-group differences in mean resource use and mean costs at each assessment point will be compared using the two-sample t-test. Statistical significance was assessed at the 5% significance level. A non-parametric bootstrap routine with bias correction for standard errors and confidence intervals will be implemented, generating 1,000 replications of the data. Estimates of standard errors surrounding mean resource use (or cost) estimates and 95% confidence intervals surrounding between-group differences in mean resource use (or costs) will be obtained from the bootstrap samples.

Cost-effectiveness analysis

Cost-effectiveness results for the base case analysis will be obtained by formulating a system of seemingly unrelated mixed-effects regressions for individual-level costs and effects, accounting for the patient-level correlations between the two and adjusting for pre-specified baseline patient characteristics. The covariates to be included in the regressions will be those selected a priori for the adjusted statistical analysis, namely age, gender and the baseline stratification factors (type of headache and geographical locality). The group sessions to which patients in the intervention as clustering variable in the intervention group and the control group will act as a separate cluster on its own. Additionally, we will control for imbalance in baseline costs and EQ-5D values between the two trial arms by including a covariate for baseline costs in the cost model and baseline health related quality of life in the QALY model, a practice that is now standard for trial-based economic evaluations (25). Estimates of the incremental costs and QALYs associated with the CHESS intervention compared with best usual care will be generated from the regressions and presented as incremental cost-effectiveness ratios (ICERs) and cost-effectiveness acceptability curves (CEACs). This accommodates sampling (or stochastic) uncertainty and

varying levels of willingness to pay for an additional QALY such as £15,000 per QALY threshold recently estimated by Claxton et al. (26) and the £20,000 to £30,000 per QALY threshold used by NICE in its technology appraisal process.(27) Heterogeneity in the trial population will be explored by formulating a net-benefit value for each patient from the observed costs and effects, and then constructing a regression model with a treatment variable and covariates such as age, gender, medication overuse and headache type where data allows us to do so. Treatment by covariate interaction terms will be included for each covariate one at a time. The magnitude and significance of the coefficients on the interaction between the covariates and the treatment variable should provide an estimate of the cost-effectiveness of the intervention by sub-group.

Additionally, due to known limitations of within-trial economic evaluations(28), we will also construct a Cohort Markov model to model beyond the parameters of the proposed within-trial cost-effectiveness of the intervention in the relevant patient population. We will inform the model with data from the trial as well as information identified from our systematic search of the literature. Long term estimates of costs and health consequences will be discounted to present values using discount rates recommended for health technology appraisal in the United Kingdom. A series of probabilistic sensitivity analyses will be undertaken to explore the implications of parameter uncertainty on the incremental cost-effectiveness ratios. All analyses will be conducted using the statistical package R (21).

Sensitivity analyses

The following sensitivity analyses will be conducted to investigate sensitivity of the base case results to:

- Utilities generated from via the SF-12/SF-6D tariff for UK (17)
- The new EQ-5D-5L tariff for England (29)
- Costs calculated from a societal perspective
- Complete case analysis as the base case cost-effectiveness analysis uses imputed attributable costs and QALYs.
- EQ-5D-5L utilities derived HIT-6 via mapping coefficients

- EQ-5D-5L utilities derived CHQLQ via mapping coefficients
- SF-6D utilities derived HIT-6 via mapping coefficients
- SF-6D utilities derived CHQLQ via mapping coefficients

Subgroup analyses

Estimates of incremental cost-effectiveness will be calculated for the following subgroup of patients.

- Medication overuse
    - Yes
    - No
- Location (Midlands versus Greater London)
- Gender (Female versus Male)
- Age group (<40years versus ≥40 years)

SECTION SEVEN: TEMPLATE TABLES AND FIGURES

Results Tables

Table 1: Completion rates for health economic outcomes

	Completion rates	
Assessment point and resource category	CHESS intervention (n=xxx)	Best usual care (n=xxx)
Baseline		
EQ-5D-5L index	xxx%	xxx%
EQ-5D-5L VAS	xxx%	xxx%
SF-12 (SF-6D) utility score	xxx%	xxx%
Hospital inpatient (admitted care)	xxx%	xxx%
Day case attendance	xxx%	xxx%
Outpatient attendance	xxx%	xxx%

Consultations (primary care) – does this need to be split by type ie GP, nurse	xxx%	xxx%
Tests and investigations (primary care)	xxx%	xxx%
Prescribed medication (primary care)		
Over the counter medication	xxx%	xxx%
Private healthcare expenditure	xxx%	xxx%
Additional costs	xxx%	xxx%
Lost income due to headache related illness	xxx%	xxx%
Time off work due to headache related illness	xxx%	xxx%
4 month assessment point		
8 month assessment point		
12 month assessment point		

Table 2: Health and social care resource utilisation during follow-up

		CHESS intervention (n=xxxx)			Best usual care (n=xxxx)			CHESS intervention versus best usual care	
Assessment point	Category	% missing	Number of visits, mean (se)	Total duration in days / minutes, mean (se)	% missing	Number of visits, mean (se)	Total duration in days/minutes, mean (se)	Total duration, mean difference (bootstrap 95% CI)[1]	P-value
Baseline	Hospital inpatient								
	Day case								
	Admitted care (overnight stay)								
	Hospital outpatient								
	Headache clinic								
	Physiotherapist								
	Occupational therapist								
	Radiology: MRI scan								
	Radiology: CT scan								
	Radiology: X-ray								
	Radiology: Ultrasound								
	Blood tests[2]								
	Accident and emergency								

	Other outpatient								
	Primary care								
	GP, surgery visit								
	GP, home visit								
	GP, telephone contact								
	Practice nurse								
	District nurse								
	Community physiotherapist								
	Occupational therapist								
	counsellor								
	Psychology/psychotherapy								
	Social worker								
	Any other contact								
4 month assessment point									
8 month assessment point									

12 month assessment point									

[1]mean difference and 95% bias corrected bootstrap confidence intervals

Table 3: Health and social care costs incurred during trial follow-up

		CHESS intervention (n=xxxx)			Best usual care (n=xxxx)			CHESS intervention versus best usual care	
Assessment point	Category	% missing	% zero costs	Mean costs (se)	% missing	% zero costs	Mean costs (se)	Mean difference, (bootstrap 95% CI)[1]	P-value
Baseline	Hospital inpatient								
	Day case								
	Admitted care (overnight stay)								
	Total inpatient costs								
	Hospital outpatient								
	Headache clinic								
	Physiotherapist								
	Occupational therapist								
	Radiology: MRI scan								
	Radiology: CT scan								

	Radiology: X-ray								
	Radiology: Ultrasound								
	Accident and emergency								
	Other outpatient								
	Total outpatient costs								
	Primary care								
	GP, surgery visit								
	GP, home visit								
	GP, telephone contact								
	Practice nurse								
	District nurse								
	Community physiotherapist								
	Occupational therapist								
	counsellor								
	Psychology/psychotherapy								
	Social worker								
	Any other contact								
	Total primary care costs								

[1]mean difference and 95% bias corrected bootstrap confidence intervals

Table 4: Private health care resource use during follow-up

		CHESS intervention(n=xxxx)			Best usual care (n=xxxx)			CHESS intervention versus best usual care	
Assessment point	Category	% missing	Number of visits, mean (se)	Total duration in days / minutes, mean (se)	% missing	Number of visits, mean (se)	Total duration in days / minutes, mean (se)	Total duration, mean difference (bootstrap 95% CI)[1]	P-value
Baseline	Over the counter medication								
	Physiotherapist								
	Occupational therapist								
	Counsellor								

	Psychologist								
	Radiology: MRI scan								
	Radiology: CT scan								
	Radiology: X-ray								
	Radiology: Ultrasound								
	Consultant service								
	Osteopath								
	Chiropractor								
	Acupuncturist								
	Homeopath								
	Other								
4 month assessment point									
8 month assessment point									
12 month assessment point									

[1]mean difference and 95% bias corrected bootstrap confidence intervals

Table 5: Private healthcare costs incurred during follow-up

		CHESS intervention (n=xxxx)			Best usual care (n=xxxx)			CHESS intervention versus best usual care	
Assessment point	Category	% missing	% zero costs	Mean costs (se)	% missing	% zero costs	Mean costs (se)	Mean cost difference, (bootstrap 95% CI)[1]	P-value
Baseline	Over the counter medication								
	Physiotherapist								
	Occupational therapist								
	Counsellor								
	Psychologist								
	Radiology: MRI scan								
	Radiology: CT scan								
	Radiology: X-ray								
	Radiology: Ultrasound								
	Consultant service								
	Osteopath								
	Chiropractor								
	Acupuncturist								
	Homeopath								

	Other								
	Total baseline costs								
4 month assessment point									
8 month assessment point									
12 month assessment point									

[1]mean difference and 95% bias corrected bootstrap confidence intervals

Table 6: Additional costs incurred during trial follow-up

		CHESS intervention (n=xxxx)			Best usual care (n=xxxx)			CHESS intervention versus best usual care	
Assessment point	Category	% missing	Number of visits, mean (se)	Total number of days, mean (se)	% missing	Number of visits, mean (se)	Total number of days, mean (se)	Mean difference, (bootstrap 95% CI)[1]	P-value
3 months post randomisation	Travel costs (e.g. bus fares)								
	Child care costs								
	Income lost								

	Cost of help with housework								
	Cost of laundry services								
	Other additional costs								
	Total additional costs								
4 month assessment point									
8 month assessment point									
12 month assessment point									

[1]mean difference and 95% bias corrected bootstrap confidence intervals for total number of days or number of contacts/visits when number of days is not relevant

Table 7: Sources of unit costs information

Category	Currency code	Unit cost	Source
Inpatients (per day of inpatient stay)			
Day case			
Admitted care			
Accident and emergency			
Out patients (per contact)			
General surgery			
ENT			
Accident and Emergency			
Pain clinic			
General Medicine			
Diabetes			
Cardiology			
Dermatology			
Breast clinic			
Neurology			
Rheumatology			
Dentist			
Eye Clinic			
Gynaecology			
Midwife			
Osteopath			
Physiotherapy			
Chiropractor			
Podiatrist			
Mental health			
Blood test			
Occupational health			
MRI Scan			
CT Scan			
X-Ray scan			
Primary and social care (cost per contact)			
Acupuncture			
Chiropractor			
Physiotherapy			
Osteopathy			

Massage			
Pharmacist			
Psychology			
Counsellor			
District nurse/ health visitor / midwife			
Practice nurse			
GP home visit			
GP surgery			
GP telephone			
Health care assistant			
Private costs			
Physiotherapy			
Psychology			

Table 8: Total economic costs

Costing perspective and list of included cost categories	CHESS intervention (n=xxxx)			Best usual care (n=xxxx)			CHESS intervention versus best usual care	
	% missing	% zero costs	Mean (SE), £	% missing	% zero costs	Mean (SE), £	Mean difference (bootstrap 95% CI), £	P-value
NHS/PSS perspective								
Intervention costs								
Follow-up costs								
Total NHS/PSS costs								
Societal perspective								
Intervention costs								
Follow-up costs (NHS/PSS)								
Follow-up costs (non-NHS/PSS)								
Total societal costs								

[1]Confidence intervals obtained by bootstrap bias corrected percentile method

Table 9: Summary of EQ5D-5L responses and scores on the visual analogue (VAS) scale

	EQ-5D dimension/ response	CHESS intervention (n=xxxx)	Best usual care (n=xxxx)	p-value[1]
Baseline	*Mobility* No problems Slight problems Moderate problems Severe problems Unable to walk Missing			
	Self-care No problems Slight problems Moderate problems Severe problems Unable to wash/dress Missing			
	Usual activities No problems Slight problems Moderate problems Severe problems Unable to do usual activities Missing			
	Pain and discomfort No problems Slight problems Moderate problems Severe problems Extreme pain and discomfort Missing			
	Anxiety and depression No problems Slight problems Moderate problems Severe problems Extremely anxious/depressed Missing			
	Visual analogue score Mean score (SE)			

	Missing			
4 months assessment point	*Mobility* No problems Slight problems Moderate problems Severe problems Unable to walk Missing			
	Self-care No problems Slight problems Moderate problems Severe problems Unable to wash/dress Missing			
	Usual activities No problems Slight problems Moderate problems Severe problems Unable to do usual activities Missing			
	Pain and discomfort No problems Slight problems Moderate problems Severe problems Extreme pain and discomfort Missing			
	Anxiety and depression No problems Slight problems Moderate problems Severe problems Extremely anxious/depressed Missing			
	Visual analogue score Mean score (SE) Missing			
8 months assessment point	*Mobility* No problems Slight problems Moderate problems			

	Severe problems Unable to walk Missing			
	Self-care No problems Slight problems Moderate problems Severe problems Unable to wash/dress Missing			
	Usual activities No problems Slight problems Moderate problems Severe problems Unable to do usual activities Missing			
	Pain and discomfort No problems Slight problems Moderate problems Severe problems Extreme pain and discomfort Missing			
	Anxiety and depression No problems Slight problems Moderate problems Severe problems Extremely anxious/depressed Missing			
	Mobility No problems Slight problems Moderate problems Severe problems Unable to walk Missing			
	Visual analogue score Mean score (SE) Missing			
12 months assessment point	*Self-care* No problems Slight problems			

	Moderate problems Severe problems Unable to wash/dress Missing			
	Usual activities No problems Slight problems Moderate problems Severe problems Unable to do usual activities Missing			
	Pain and discomfort No problems Slight problems Moderate problems Severe problems Extreme pain and discomfort Missing			
	Anxiety and depression No problems Slight problems Moderate problems Severe problems Extremely anxious/depressed Missing			
	Visual analogue score Mean score (SE) Missing			

[1]P-values were generated from chi-squared tests for differences in sub-optimal levels of function for each dimension where responses indicating no functional impairment were categorised as optimal and responses indicating any functional impairment were categorised as sub-optimal.

Table 10: SF-12 v2 responses

Assessment point	Response	CHESS intervention (n=xxxx)	Best usual care (N=xxxx)	P-value[1]
Baseline	*General health* Excellent Very good Good Fair Poor Missing			
	Moderate activities Yes, limited a lot			

	Yes, limited a little No, not limited at all Missing			
	Climbing stairs Yes, limited a lot Yes, limited a little No, not limited at all Missing			
	Accomplished less physically All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Limited physically All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Did less Work emotional All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Accomplished less emotionally All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Pain Not at all A little bit Moderately Quite a bit Extremely Missing			
	Calm All the time Most of the time A good bit of the time Some of the time A little bit of the time			

	None of the time Missing			
	Energy All of the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Feeling down hearted All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Social activities All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
4 months post randomisation	*General health* Excellent Very good Good Fair Poor Missing			
	Moderate activities Yes, limited a lot Yes, limited a little No, not limited at all Missing			
	Climbing stairs Yes, limited a lot Yes, limited a little No, not limited at all Missing			
	Accomplished less physically All of the time Most of the time Some of the time A little of the time None of the time Missing			

	Limited physically All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Did less Work emotional All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Accomplished less emotionally All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Pain Not at all A little bit Moderately Quite a bit Extremely Missing			
	Calm All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Energy All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Feeling down hearted All the time Most of the time A good bit of the time Some of the time			

	A little bit of the time None of the time Missing			
	Social activities All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
8 month post-randomisation	*General health* Excellent Very good Good Fair Poor Missing			
	Moderate activities Yes, limited a lot Yes, limited a little No, not limited at all Missing			
	Climbing stairs Yes, limited a lot Yes, limited a little No, not limited at all Missing			
	Accomplished less physically All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Limited physically All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Did less Work emotional All of the time Most of the time Some of the time A little of the time None of the time Missing			

	Accomplished less emotionally All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Pain Not at all A little bit Moderately Quite a bit Extremely Missing			
	Calm All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Energy All of the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Feeling down hearted All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Social activities All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
12 months post randomisation	*General health* Excellent			

	Very good Good Fair Poor Missing			
	Moderate activities Yes, limited a lot Yes, limited a little No, not limited at all Missing			
	Climbing stairs Yes, limited a lot Yes, limited a little No, not limited at all Missing			
	Accomplished less physically All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Limited physically All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Did less Work emotional All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Accomplished less emotionally All of the time Most of the time Some of the time A little of the time None of the time Missing			
	Pain Not at all A little bit Moderately Quite a bit Extremely			

	Missing			
	Calm All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Energy All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Feeling down hearted All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			
	Social activities All the time Most of the time A good bit of the time Some of the time A little bit of the time None of the time Missing			

Table 11: Summary of health-related quality of life (utility) scores generated from EQ-5D-5L and SF-12 v2 instruments

Outcomes	CHESS intervention			Best usual care			CHESS intervention versus best usual care	
	N	% missing	Mean (SE)	N	% missing	Mean (SE)	Mean difference (95% CI)	P-value
EQ-5D-5L to 3L cross walk[1]								
Baseline	xxxx			xxxx				
4 months	xxxx			xxxx				
8 months	xxxx			xxxx				
12 months	xxxx			xxxx				
EQ-5D-5L (new UK tariff)[2]								
Baseline	xxxx			xxxx				
4 months	xxxx			xxxx				
8 months	xxxx			xxxx				
12 months	xxxx			xxxx				
SF-12 (SF-6D UK tariff)								
Baseline	xxxx			xxxx				
4 months	xxxx			xxxx				
8 months	xxxx			xxxx				
12 months	xxxx			xxxx				
EQ-5D-5L VAS								
Baseline	xxxx			xxxx				
4 months	xxxx			xxxx				
8 months	xxxx			xxxx				
12 months	xxxx			xxxx				

[1]The EQ-5D-5L cross-walk utility values were derived using the interim 5L to 3L cross-walk tariffs for the UK (15)

[2]New EQ-5D-5L value set for England (13)

Table 12: Unadjusted estimates of Quality-Adjusted Life Years (QALYs) accrued over 12 months of follow-up

Outcome measure	CHESS intervention			Best usual care plus relaxation			CHESS intervention versus best usual care	
	N	% missing	Mean (SE)	N	% missing	Mean (SE)	Mean difference (95% CI)	P-value
EQ-5D-5L cross-walk tariff	xxxx			xxxx				
EQ-5D-5L (New 5L tariff for England)	xxxx			xxxx				
SF-12 (SF-6D tariff)	xxxx			xxxx				

Table 13: Cost-effectiveness of the CHESS intervention compared with best usual care based on the within-trial economic analysis

	Cost-effectiveness outcomes				Probability CHESS intervention is cost-effective at cost-effectiveness threshold of		
Description	Mean incremental costs (95% CI), £	Mean incremental QALYs (95% CI)	ICER[4]		£13,000 per QALY	£20,000 per QALY	£30,000 per QALY
Base case analysis[1]							
Sensitivity analyses Unadjusted analysis Complete case analysis Restricted to trial participants who did not participate in process evaluation interviews SF-12/SF-6D EQ-5D utilities derived HIT-6 via mapping coefficients EQ-5D utilities derived CHQLQ via mapping coefficients SF-6D utilities derived HIT-6 via mapping coefficients SF-6D utilities derived CHQLQ via mapping coefficients							
Sub-group analyses Headache type Chronic tension type headache Probable chronic migraine Definitive chronic migraine							
Headache type with medication overuse Chronic tension type headache Probable chronic migraine							

Definitive chronic migraine							
Headache type without medication overuse Chronic tension type headache Probable chronic migraine without Definitive chronic migraine without							
Geographical location Midlands Greater London							
Sex Female Male							
Age group <40years ≥40 years							

ICER = Incremental cost-effectiveness ratio; CI = confidence interval

1Adjusted for treatment allocation, age, gender, baseline stratification factors (type of headache and geographical locality), baseline health-related quality of life (QALY model) and baseline costs (cost model)

REFERENCES

1. Ramsey S, Willke R, Briggs A, Brown R, Buxton M, Chawla A. Good Research Practices for Cost-Effectiveness Analysis Alongside Clinical Trials: The ISPOR RCT-CEA Task Force Report. Value in Health. 2005;8:521-33.

2. NICE. Guide to the methods of technology appraisal. NICE [Internet]. 2013 12 15. Available from: https://www.nice.org.uk/process/pmg9/resources/guide-to-the-methods-of-technology-appraisal-2013-pdf-2007975843781.

3. Sanders GD, Neumann PJ, Basu A, et al. Recommendations for conduct, methodological practices, and reporting of cost-effectiveness analyses: Second panel on cost-effectiveness in health and medicine. JAMA. 2016;316(10):1093-103.

4. Husereau D, Drummond M, Petrou S, Carswell C, Moher D, Greenberg D, et al. Consolidated Health Economic Evaluation Reporting Standards (CHEERS) statement. BMJ : British Medical Journal. 2013;346.

5. Curtis L. Unit costs of health and social care. Canterbury, UK: University of Kent; 2011.

6. Glick HA, Doshi JA, Sonnad SS, Polsky D. Economic Evaluation in Clinical Trials. 2nd ed. Gray A, Briggs A, editors. Oxford: Oxford University Press; 2015.

7. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality of life research : an international journal of quality of life aspects of treatment, care and rehabilitation. 2011;20(10):1727-36.

8. Ware JE. How to Score Version 2 of the SF-12v2® Health Survey (With a Supplement Documenting SF-12® Health Survey) Lincoln, RI.: QualityMetric Inc; 2002.

9. Dolan P, Roberts J. Modelling valuations for Eq-5d health states: an alternative model using differences in valuations. Med Care. 2002;40(5):442-6.

10. Brazier J, Czoski-Murray C, Roberts J, Brown M, Symonds T, Kelleher C. Estimation of a preference-based index from a condition-specific measure: the King's Health Questionnaire. Medical decision making : an international journal of the Society for Medical Decision Making. 2008;28(1):113-26.

11. EuroQol G. EuroQol--a new facility for the measurement of health-related quality of life. Health Policy. 1990;16(3):199-208.

12. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Quality in Life Research. 2011;20(10):1727-36.

13. 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. 2017:n/a-n/a.

14. National Istitute for Health and Care Excellence. Position statement on use of the EQ-5D-5L valuation set. Available from

https://www.nice.org.uk/Media/Default/About/what-we-do/NICE-guidance/NICE-technology-appraisal-guidance/eq5d5l_nice_position_statement.pdf. Accessed on 28 September 2017. 2017.

15. van Hout B, Janssen MF, Feng Y-S, Kohlmann T, Busschbach J, Golicki D, et al. Interim Scoring for the EQ-5D-5L: Mapping the EQ-5D-5L to EQ-5D-3L Value Sets. Value in Health. 2012;15(5):708-15.

16. Kind P, Dolan P, Gudex C, Williams A. Variations in population health status: results from a United Kingdom national questionnaire survey. Brit Med J. 1998;316(7133):736 - 41.

17. Brazier JE, Roberts J. The Estimation of a Preference-Based Measure of Health From the SF-12. Medical Care. 2004;42(9):851-9.

18. Gomes M, Grieve R, Nixon R, Carpenter J, Thompson SG. Developing appropriate methods for cost-effectiveness analysis of cluster randomized trials. Medical Decision Making. 2012;32(2):350-61.

19. Ramsey SD, Willke RJ, Glick H, Reed SD, Augustovski F, Jonsson B, et al. Cost-Effectiveness Analysis Alongside Clinical Trials II 2014;An ISPOR Good Research Practices Task Force Report. Value in Health.18(2):161-72.

20. Van Buuren S, Groothuis-Oudshoorn K. Mice: Multivariate Imputation by Chained Equations. Journal of Statistical Software, 45, 1-67. . 2011.

21. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/. 2017.

22. Sterne JA, White IR, Carlin JB, Spratt M, Royston P, Kenward MG. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. BMJ. 2009;338:2393.

23. Faria R, Gomes M, Epstein D. A Guide to Handling Missing Data in Cost-Effectiveness Analysis Conducted Within Randomised Controlled Trials. PharmacoEconomics. 2014;32:1157-70.

24. Faria R, Gomes M, Epstein D, White IR. A Guide to Handling Missing Data in Cost-Effectiveness Analysis Conducted Within Randomised Controlled Trials. PharmacoEconomics. 2014;32(12):1157-70.

25. Manca A, Hawkins N, Sculpher MJ. Estimating mean QALYs in trial-based cost-effectiveness analysis: the importance of controlling for baseline utility. Health Econ. 2005;14(5):487-96.

26. Claxton K, Martin S, Soares M, Rice N, Spackman E, Hinde S, et al. Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold. Health technology assessment (Winchester, England). 2015;19(14):1-503, v-vi.

27. National Istitute for Health and Care Excellence. Guide to the methods of technology appraisal 2013. Avaialable from http://publications.nice.org.uk/pmg9. 2013.

28. Sculpher MJ, Claxton K, Drummond M, McCabe C. Whither trial-based economic evaluation for health care decision making? Health Econ. 2006;15(7):677-87.

29. Devlin N, Shah K, Feng F, Mulhern B, van Hout B. Valuing Health-Related Quality of Life: An EQ-5D-5L Value Set for England. Available from https://www.ohe.org/publications/valuing-health-related-quality-life-eq-5d-5l-value-set-england. 2016.