**Maintaining Musculoskeletal Health using a behavioural therapy approach: a population-based randomised controlled trial (The MAmMOTH study)**

Gary J Macfarlane MD(Hons) (1), Marcus Beasley MSc (1), Neil W Scott PhD (2), Huey Chong PhD (3), Paul McNamee PhD (3), John McBeth PhD (4), Neil Basu PhD (5), Philip C Hannaford MD (6) , Gareth T Jones PhD (1), Phil Keeley PhD (7), Gordon Prescott PhD (8), Karina Lovell PhD (9)

(1)Aberdeen Centre for Arthritis and Musculoskeletal Health (Epidemiology Group), University of Aberdeen (2) Medical Statistics Team, University of Aberdeen (3) Health Economics Research Unit, University of Aberdeen (4) Versus Arthritis Centre for Epidemiology, University of Manchester (5) Institute of Infection, Immunity and Inflammation, University of Glasgow (6) Academic Primary Care, University of Aberdeen (7) School of Nursing and Midwifery, University of Keele (8) Lancashire Clinical Trials Unit, University of Central Lancashire (9) Division of Nursing, Midwifery and Social Work, University of Manchester

Author for Correspondence:

Professor Gary J Macfarlane

University of Aberdeen

King's College

Aberdeen AB24 3FX

E-mail: g.j.macfarlane@abdn.ac.uk

Twitter: @AberdeenEpi

**Abstract**

Objective: Cognitive Behaviour Therapy (CBT) has been shown to be effective in the management of Chronic Widespread Pain (CWP); we now test whether it can *prevent* onset among adults at high risk.

Methods: A population-based randomised controlled prevention trial, with recruitment through UK general practices. A mailed screening questionnaire identified adults at high risk of CWP. Participants received either usual care (UC) or a short course of telephone CBT. The primary outcome was CWP onset at 12 months assessed by mailed questionnaire. There were seven secondary outcomes including quality of life (EQ-5D-5L) used as part of a health economic assessment.

Results: 996 participants were randomised and included in the intention-to-treat analysis of which 825 provided primary outcome data. The median age of participants was 59 years; 59% were female. At 12 months there was no difference in the onset of CWP (tCBT: 18.0% v. UC: 17.5%; OR 1.05; 95% CI 0.75-1.48). Participants who received tCBT were more likely to report better quality of life (EQ-5D-5L utility score mean difference 0.024 (95% CI 0.009-0.040)); and had 0.023 (95% CI 0.007-0.039) more QALYs at an additional cost of £42.30 (95% CI −451.19-597.90), yielding an incremental cost effectiveness ratio of £1,828. Most secondary outcomes showed significant benefit for the intervention.

Conclusions: A short course of tCBT does not prevent onset of CWP in adults at high risk, but did improve quality of life and was cost-effective. A low-cost, short duration, intervention benefits persons at risk of CWP.

**Key Messages:**

What is already known about this subject?

- Cognitive Behaviour Therapy (CBT) has demonstrated long-term effectiveness in managing chronic widespread pain, the characteristic symptom of fibromyalgia
- It improves patient global assessment of change and quality of life

What does this study add?

- A short course of telephone CBT in persons evaluated at high risk of developing CWP does not change onset of CWP but does result in a wide range of health benefits including improved quality of life.

How might this impact on clinical practice or future developments?

- CBT derives benefit for a wider group of people with pain than previously established and in relation to this wider group is highly cost-effective

## Introduction

Chronic widespread pain (CWP) is common, with an estimated population prevalence of 10.6% (95% confidence interval (CI) 8.6-12.9) (1) and is the key feature of fibromyalgia which is the second most common reason (after osteoarthritis) for referral to a rheumatologist (2). Chronic widespread pain and fibromyalgia result in a substantial impact on health-related quality of life (3) even in comparison to other musculoskeletal disorders (4).

The road to diagnosis is often tortuous and can take many years. Using general practitioner records in the United Kingdom, Hughes et al (5) noted that people diagnosed with fibromyalgia had higher rates of primary care visits (average 25 visits/year), prescriptions (11/year), and testing from at least 10 years prior to diagnosis, in comparison to matched persons without such a diagnosis (12 visits/year and 4.5 prescriptions/year). Current European guidelines emphasise the primary role of non-pharmacological therapies for fibromyalgia (6). Evidence in relation to musculoskeletal pain generally, is that the longer the duration of symptoms, the less likely they are to improve including with specific interventions (7).

A Versus Arthritis "Research roadmap for pain" produced by scientists, clinicians and patients identified preventing future musculoskeletal pain as one of four main priorities (8). Further recognising its importance, the International Association for the Study of Pain nominated 2020 as "The Global Year for the Prevention of Pain". Despite this, we are not aware of any large-scale trials which have tested approaches to the future prevention of pain.

We have previously shown, in a randomised controlled trial, short and long-term effectiveness of a course of Cognitive Behaviour Therapy delivered by telephone (tCBT) for CWP, compared to usual care (9.10). These results are consistent with a meta-analysis of 29 trials involving 2509 participants and comparing CBT (across all modes of delivery) with control interventions for the management of fibromyalgia, which found high quality evidence for improving pain, reducing disability, negative mood and fatigue (11).

We have developed, validated and refined a statistical model which identifies people at high risk for the future development of CWP (12,13). On the basis of reporting somatic symptoms, sleep problems and aspects of illness behaviour, those classified as "high risk" have around 1 in 4 chance of reporting CWP one year later. Therefore, building upon the evidence for the use of tCBT in the management of CWP and the ability to identify those with risk factors for its development, we undertook a trial to test whether tCBT can reduce CWP onset amongst those at high risk.

## Methods

*Study Design*

We conducted a randomised controlled parallel prevention trial, recruiting through a population-based sampling frame, in three health boards within the United Kingdom (NHS Grampian, NHS Greater Glasgow and Clyde, NHS

Highland), the protocol for which has been previously published (14). Recruitment was through sixteen general practices. Ethical approval was obtained from Cornwall and Plymouth Research Ethics Committee Reference 16/SW/0019.

*Participants*

A short screening questionnaire, to determine eligibility for the trial, was mailed to persons aged 25 years and over registered at participating general practices in the study area. Respondents eligible for the trial were those assessed as at high risk of developing CWP, namely that they reported pain which did not satisfy the definition of CWP used in the 1990 American College of Rheumatology criteria for fibromyalgia (namely axial and contra-lateral body pain present for at least 3 months), and hereafter referred to as "ACR criteria" (15), and satisfied at least two of the following a) a score >4 on the Illness Behaviour Subscale of the Illness Attitudes Scale (16) b) a score >2 on the Somatic Symptom Scale (SSS) score (but excluding items on pain) (17) c) a score >= 4 on the Sleep Problem Scale (18). In order to ensure that in the event of the trial showing benefit there was a relevant clinical population to which the intervention could be applied, we added to the risk models we had developed the requirement that persons had consulted to primary care within the previous six months or reported consulting a doctor frequently. Respondents were not eligible to take part if they had a medical condition which would make the proposed intervention unsuitable (e.g. lacked cognitive ability).

*Randomisation*

Potentially eligible participants were contacted by post with information about the study, and subsequently by a study researcher by telephone to confirm their willingness to take part and provide informed consent. Participants were allocated into groups using a computer randomisation program (1:1 allocation ratio) , stratified in blocks by two factors a) the number of non-pain "high-risk" factors they reported (2 or 3) since this is related to the risk of CWP onset, and b) the general practice at which they were registered.

*Procedures*

The tCBT intervention consisted of an initial assessment (45–60 min), 6 weekly sessions (each 30–45 min) over six weeks, and then booster sessions at 3 and 6 months. The intervention was delivered by therapists trained for the study and accredited by the British Association for Behaviour and Cognitive Psychotherapies. Participants were supported by a self-management manual. The therapist conducted an assessment for problem identification, and they developed with each participant a shared formulation of the current health problem. The sessions involved education about musculoskeletal pain, somatic symptoms and specific techniques such as pacing of activity, behavioural activation, diary keeping, identifying and challenging negative and unhelpful thinking patterns and the development of a longer-term management plan. Participants would record in the manuals agreed goals for the therapist and patient to work towards, and some activities to complete between sessions. Therapists delivering the intervention received a 2-day training programme conducted by the investigators. Therapists were supervised every two weeks (by investigators KL and PK) throughout the delivery of the intervention. The number of telephone consultations conducted was recorded,

although the therapist and participant could jointly agree that no further sessions were required before all planned sessions had been completed.

The group allocated to usual care received no additional intervention, reflecting the fact there is no specific intervention provided to patients currently for the prevention of CWP. There was no restriction on what this care could involve.

Follow-up questionnaires were mailed to participants at 3, 12 and 24 months after the treatment start date (for participants in the active treatment group) or dummy treatment start date (for those in usual care). The dummy treatment start date for a participant randomised to UC was determined by the treatment start date of the last participant to be randomised to receive active treatment. At 3 and 12 months, participants who did not return their questionnaire were telephoned to ask them to complete and return it, while at 24 months the follow-up call also offered the option of completing a shortened version by telephone.

*Outcomes*

The principal outcome time was at 12 months follow-up and the primary outcome was ACR criteria for CWP. Secondary outcomes were: Global Impression of Change, Illness Behaviour Subscale of the Illness Attitudes Scale (16), the SSS (excluding items on pain) (17), the Sleep Problem Scale (18), the presence of pain over the past month, Widespread Pain Index (WPI) and Symptoms Severity Scale (SSS) of the 2010 (revised) criteria for fibromyalgia (19), psychological distress measured using by the General Health Questionnaire (GHQ) (20), Chalder Fatigue Scale (21), quality of life (EQ-5D-5L) (22) and capability (ICECAP-A)(23). Further details of secondary outcome (including coding) are given in the Supplementary Text file.

*Statistical Analysis*

All analyses were undertaken using STATA version 15. The *a priori* target sample size was 946 participants, which would provide 90% power to detect a group difference of 9% (21% vs 12%) in the percentage of participants with chronic widespread pain (CWP) at 12 months follow-up, assuming a 5% significance level and an 80% response rate.

Where there was missing data within a scale score, we followed standard procedures (where available) as to if and how the missing values could be imputed. The analysis of the primary outcome used a binary logistic regression model with results expressed as an odds ratio (OR) with 95% CI. Secondary outcomes were analysed using linear, binary logistic, ordinal logistic, or Poisson regression models for continuous, binary, ordinal, and count variables respectively. Model results were reported using mean differences, odds ratios or incident rate ratios (IRRs) as appropriate. Except for EQ-5D-5L, mean differences less than zero and ORs/IRRs less than one favour the treatment group. All models were adjusted (adj) for the number of non-pain risk factors on screening (two or three), age (years), gender, general practice (random effect) and baseline score of the outcome measure (where applicable). The primary analysis was by intention to treat – i.e. participants were analysed according to randomised group regardless of the number of sessions received. Separate analyses were performed for each time point (3, 12 and 24 months). For the primary outcome a p-value less

than 0.05 was regarded as statistically significant; for secondary outcomes p<0.01 was used. Additional sensitivity analyses were conducted for the primary outcome only and are detailed in the supplementary text file.

*Health Economic Analysis*

Health service resource use over 24 months was assessed using responses from self-reported questionnaires. Participants were asked to recall their usage for the previous 4 week period at each follow-up. Resource use was then valued using published UK sources - NHS Reference Cost and the Personal and Social Service Research Unit (PSSRU) for NHS primary and secondary care and published literature for care obtained from private providers (24). The unit costs used for the valuation of health service resource use are reported in Supplementary Material Table S1. The intervention cost was based on the actual number and duration of telephone calls per participant ("direct time"), plus time spent on training and supervision. An allowance for indirect time spent was also included and this was based on an assumed ratio of 1:1 between time spent on participant contact and other activities conducted by therapist. Training costs were estimated using the time spent in training by trainers and trainees (tCBT therapists). A fortnightly supervision cost was estimated by assuming 30 sessions per therapist (30 minutes per session) were provided. Costs were expressed in 2017/18 prices. Health utility scores were assigned based on responses to the EQ-5D-5L at each follow-up, and these were converted using the 'crosswalk' procedure to EQ-5D-3L (25). There is currently no consensus on the preferred EQ-5D-5L tariff for use in economic evaluation, althoughthe National Institute for Health and Care Excellence (NICE) recommend the use of the 'crosswalk' procedure (a validated mapping function) to derive health utility scores for the EQ-5D-5L from the EQ-5D-3L tariff [https://www.nice.org.uk/about/what-we-do/our-programmes/nice-guidance/technology-appraisal-guidance/eq-5d-5l accessed 20 November 2020]. These utility scores were used to estimate quality-adjusted life years (QALYs) over the 24-month using the Area Under the Curve (AUC) method (26). Costs and QALYs incurred beyond 12 months were discounted at the rate of 3.5% per annum.

The within-trial economic analysis was conducted over 24 months from a UK National Health Service (NHS) cost perspective. To estimate the differences in mean costs and QALYs between groups, generalised linear models with adjustment for minimisation factors, baseline cost and baseline utility score were performed. A γ family with log link function and a Poisson family with power 0.5 link function were specified for the cost and QALY data, respectively. Missing data were addressed using multiple imputation by chained equations (MICE). Variance surrounding the incremental costs and QALYs was characterised using non-bootstrapping (500 iterations), with MICE (m=5) nested within the bootstrap loops (27). Cost-effectiveness acceptability curves (CEACs) were constructed, using 500 replications of each incremental cost-effectiveness ratio (ICER) and the net monetary benefit framework, to determine the probability of the alternative interventions being considered cost-effective at different willingness to pay (WTP) per QALY (£20,000-£30,000 per QALY was used are commonly applied ceiling ratios in the UK). Several sensitivity analyses were performed to explore the impact on the results of uncertainty in estimates made – (i) using complete cases of costs and QALYS, (ii) including private care costs, (iii) using alternative tCBT costing methodology (actual trial expenses incurred by therapists and the cost of a complete tCBT course) and (iv) using ICECAP tariff as the measure of effectiveness.

The trial was registered with clinicaltrials.gov (NCT02668003); and was evaluated by the Trial Steering Committee as not requiring a Data Monitoring Committee.

**Results**

Of 61257 screening questionnaires sent, between Apr 4, 2016 and Nov 4, 2016 to patients registered at 16 general practices, 18035 completed questionnaires were returned. From those returning a completed questionnaire, 2406 were identified as potentially eligible and sent invitations to take part in the trial. 1002 participants were recruited to the trial and randomised, 501 to tCBT and 501 to UC, between May 2016 - March 2017. Six participants were subsequently determined to be ineligible for the trial and were excluded from analyses (see Trial Profile: Figure 1) leaving a final study size of 500 and 496 in the tCBT and UC arms respectively. At the 3, 12- and 24-month follow-up there were 823, 825 and 853 respondents who provided primary outcome data, respectively. Most participants (51%) came from the lowest two quintiles of deprivation, while 18% came from the two most deprived quintiles.

Participants at the time of recruitment had a median age of 59 years (inter-quartile range (IQR) 48-69), 59% were female, and 52% were working full-time or part-time (Table 1). The median EQ-5D utility score was 0.74 IQR (0.65–0.80). The vast majority satisfied only two of the non-pain criteria for eligibility, nearly always on the basis of a high score on the illness behaviour subscale of the Illness attitudes scale and having sleep problems. Only 6% of the study sample satisfied the somatic symptoms criterion. The tCBT and UC groups were well matched in terms of the measured health-related factors.

Results for all outcome measures at the primary time point (12 months) are shown in Table 2. The corresponding results at 3 and 24 months are shown in Supplementary Material (Tables S2-S3). Table 3 provides a summary of all primary and secondary outcomes at all time points and shows adjusted and unadjusted effect sizes.

*Primary outcome*

At the 12 month time point similar percentages in the tCBT and UC groups reported having CWP (tCBT: 69/384 (18.0%), UC: 77/441 (17.5%); adj OR 1.05; 95% CI: 0.75-1.48; difference in percentages: adj 0.73, 95% CI: -4.15-5.61) (Tables 2-3). Very similar results were obtained at 3 months (17.9% v. 16.9%; adj OR: 1.08; 95% CI: 0.74-1.58) and 24 months (19.6% v. 22.3%; adj OR: 0.85; 95% CI: 0.68-1.07) (Table S2-S3, Table 3). There was no difference in the interpretation when examining unadjusted results, per protocol results or the analyses using multiple imputation (Table 3). The GEE model, incorporating data from all three time points, also showed no evidence of a difference (adj OR: 1.00; 95% CI: 0.96 -1.04; p=0.91).

*Secondary outcomes*

At 12 months, those randomised to tCBT were more likely to perceive their health to be improved (adj OR (OLR): 0.51, 95% CI: 0.39-0.67) and to report better quality of life (EQ-5D-5L utility scores) (adj mean diff: 0.024, 95% CI: 0.009-0.040) (Tables 2-3). While those who received tCBT had lower illness behaviour (adj mean difference (diff): -0.81; 95% CI: -1.54--0.09) and sleep problem scores (adj mean diff: -0.95; 95% CI: -1.48--0.42), but there was no significant difference in relation to somatic symptoms (adj OR): 0.86; 95% CI: 0.71-1.04). Participants randomised to tCBT had improved distress (GHQ scores) (adj OR: 0.65, 95% CI: 0.50-0.86) and lower levels of fatigue (Chalder scale scores) (adj mean diff: -1.02, 95% CI: -1.63--0.42). There was no evidence of a difference for ICECAP-A tariffs (adj OR (OLR): 1.39, 95% CI: 0.94-2.04; p=0.10). In relation to the components of criteria for fibromyalgia, they had lower scores on the WPI (adj IRR: 0.88; 95% CI: 0.80-0.98) and SSS (adj mean diff: -0.52, 95% CI: -0.75--0.28). Of these receiving tCBT, 3.8% met fibromyalgia research criteria at follow-up (in comparison to 6.0% amongst those receiving UC).

*Outcomes across timepoints*

Sensitivity analyses, unadjusted results and findings at 3- and 24-month time points generally yielded similar observations as those for 12 months (Tables 3, S1-S2). There was consistently no effect on the primary outcome. The strongest and most consistent effects were on patient global assessment of change – which showed large and consistent effects across all time points. There were also clear effects of the intervention (in comparison to UC) across all time points with respect to improvement in levels of fatigue and psychological distress. Quality of life was better in the intervention group from twelve months onwards. There was only one serious adverse event reported, it was in the intervention group but unrelated to the intervention.

*Health economic analysis*

The unadjusted health service resource use and costs per participant are summarised in Table S4. Participants randomised to tCBT group had an average time of 139 minutes of direct contact with therapists over the 6-month t-CBT course, and the average tCBT cost was £270.19 per participant. Compared to the usual care group, NHS primary and secondary care costs were lower amongst tCBT group, and private care costs higher. All cost-effectiveness analyses showed that tCBT was associated with an increase in health service costs and an increase in QALYs (Table 4). The primary analysis generated a mean of 0.023 (95% CI 0.007-0.039) more QALYs per participant at an additional cost of £42.30 (95%CI −£451.19-£597.90), yielding an incremental cost effectiveness ratio (ICER) of £1,828. Based on the results of the non-parametric bootstrap, tCBT was found to have a 91.6% chance of being the preferred strategy at a ceiling ratio of £20,000 per QALY gained (Figure 2). Sensitivity analyses showed that this finding was robust to changes in study perspective, inclusion of complete cases only and different assumptions relating to delivery of the intervention in terms of tCBT staff time (Figure S1 a-d).

**Discussion**

A short course of tCBT amongst persons at high risk did not change the proportion of people developing CWP (compared to UC). Those receiving the active intervention were more likely to perceive their health as having improved and report better quality of life as well as lower levels of fatigue and psychological distress. The intervention was highly cost effective in terms of incremental cost per QALY gained.

Undertaking a primary prevention study presents different challenges to undertaking a treatment study. Most people eligible for the trial probably would not have known what CWP is, nor that they were at high risk of its development. Thus, the intervention was described as "maintaining musculoskeletal health" and introduced in the context of participants having reported pain and other symptoms. Although a set number of sessions for the intervention was planned, it was agreed that at any point the intervention could be stopped with mutual agreement between therapist and participant; with the intervention considered completed. Amongst participants, 329 (66%) were considered to be completers i.e. had the assessment session and either had at least two completed treatment sessions (n= 297) or had the assessment session and up to one treatment session with mutual agreement that the intervention was complete (n=32). Of those classed as "non-completers", 97 had no assessment while 75 had an assessment and up to one treatment session.

Why did the trial clearly not change the likelihood of CWP onset while showing positive effects for a range of secondary outcomes (including quality of life)? Firstly, it may be that CBT is not effective in relation to preventing CWP onset. We know that there is a large body of evidence that CBT (including tCBT) is effective in relation to managing CWP, and also for managing some of the symptoms which characterised people at high risk, but it may not be effective at improving the pain in CWP. Our previous trial using CBT in the management of CWP while showing large improvement in patient perception of their condition and in quality of life, did not demonstrate any benefit in terms of the Chronic Pain Grade (10). Secondly our risk model may not be the causal model. A change in hypothesised risk factors would only effect a change in outcome if the relationship was causal. This suggests that it would be beneficial to explore, amongst those at risk, what is the underlying causal mechanism. Altered hypothalamic-pituitary-adrenal (HPA) axis function is one possible underlying causal mechanism which has been investigated (28). Thirdly, it is understood that there are life-course influences, specifically early life factors, on the development of CWP (29), so it could be that intervening across the adult age range is too late to be effecting a change by means of a short-term intervention. Fourthly it may be that CWP was a poor choice as the primary outcome. There is evidence that people with CWP can move in and out of meeting criteria (30) and indeed it may be that we have identified people who commonly experience CWP but recruited them at a time when they did not meet criteria - and the interpretation would be that the intervention did not move participants off that trajectory. Recent data from a longitudinal study in Norway has shown that the transition, amongst people with pain, to CWP did not represent a clinically significant change in state (31).

It is already known that CBT is effective in the management of fibromyalgia (11) and this study provides evidence that a wider range of patients may benefit in terms of quality of life. In total 54.5% of the intervention group considered

their health had improved (between a little and very much) compared to 36.9% of the usual care group, as well as improvements in fatigue, distress and changes in response to symptoms. The incremental cost per QALY gained of £1828 (which was robust to different assumptions modelled in various sensitivity analyses) means that this intervention is highly likely to be cost effective at the limit which the National Institute of Health and Clinical Excellence (NICE) in the UK, is willing to pay. In terms of delivering behavioural therapies, it has long been recognised that there is a shortage of clinical psychologists in the United Kingdom. It is not necessary to have such persons delivering behavioural therapy to all such patients even where cognitive behavioural therapy is identified as appropriate. In this study, the intervention was delivered by therapists accredited by the British Association for Behaviour and Cognitive Psychotherapies (BABCP). At a minimum this requires a Bachelor of Science degree and a two-year course leading to a postgraduate diploma in cognitive behaviour psychotherapies (CBP). Further there has been a considerable amount of research in terms of internet-based therapies. The potential advantage of such a self-directed approach is that it requires less input by the therapist (usually somewhere between 1-15 mins/week). Further, a meta-analysis of 20 studies involving 1460 participants showed that internet delivered CBT was effective in the treatment of insomnia (32) while a meta-analysis of 20 studies involving 1418 participants comparing face-to-face and internet delivered CBT for psychiatric and somatic symptoms found that "there was no evidence to conclude that they were not equivalent" (33). Studies have also examined training members of the care team (usually nurses) to deliver behavioural therapy in terms of making any service for chronic pain sustainable, and these have been shown to be effective (34). Thus we need to consider different professionals and ways of delivering CBT, particularly if we widen the group eligible to receive it, and there is no doubt that the large changes to how health services are delivered, caused by COVID-19, will only accelerate moves to the greater use of remote delivery of care

In summary, this trial has shown that a short course of tCBT does not prevent the onset of CWP in adults assessed as being at high risk. It did however positively change most other health indicators measured, including quality of life, and was highly cost effective. It demonstrates that a low-cost short duration intervention benefits a wider range of people with musculoskeletal symptoms than previously considered.

**Contributors**

GJM was CI and MB study co-ordinator. KL and PK were responsible for designing and overseeing the delivery of the intervention. GP was the trial statistician and designed the analysis plan, and this role was latterly taken over by NS who conducted the analysis. GM was director of the trial unit affiliated to the study. PM was responsible for designing the health economic analysis which was undertaken by HC. GJM drafted the manuscript with input from MB, HC, PM and NS. All authors contributed important intellectual content to trial design and execution, and commented on drafts.

**Declaration of Interests**

We declare no competing interests

**Data availability statement**

There is an application process by which researchers may request to access data in this manuscript. In principle we are willing to share de-identified data

**Acknowledgements**

The study was funded by Arthritis Research UK (now Versus Arthritis) grant number: 20748. Costs for delivery of the intervention were provided by NHS Grampian, NHS Greater Glasgow and Clyde, and NHS Highland. The funder of the study had no role in the study design, data collection, data analysis, data interpretation, or writing of the report. We acknowledge the contribution of the trial steering committee to the successful conduct of the study. The members were Professor Ernest Choy (Cardiff University), Professor Tamar Pincus (Royal Holloway, University of London) and Gordon Taylor (Bath University). We thank Brian Taylor and Mark Forrest from the Centre for Healthcare Randomised Trials (CHaRT) at the University of Aberdeen for their technical assistance and Professor Graeme MacLennan, Director of CHaRT, for methodological input. Professor John Norrie (originally University of Aberdeen now University of Edinburgh) and Dr. Majid Artus (originally Keele University, now the Osmaston surgery, Derbyshire) were study investigators at the time of grant award but subsequently left the study. We thank Kathy Longley (a representative of Fibromyalgia Action UK) for her input to the grant application and the project as well as from members of the public on the University of Aberdeen College of Life Sciences and Medicine Research Interest Group. The prioritisation of "Prevention of chronic pain" arose from a 2012 meeting of the Arthritis Research UK Clinical Study Group in Pain to which patients contributed.

**Figure Legends**

Figure 1: Trial profile

Figure 2: Cost-effectiveness plane and cost-effectiveness acceptability curve between groups (primary analysis using imputed dataset, NHS perspective). Cost-effectiveness planes were based on 500 bootstrap cost-effect pairs (adjusted for age, gender, number of risk factors present, employment status, centre, baseline EQ-5D health utility score and baseline cost). QALY, quality adjusted life year; tCBT, telephone-delivered cognitive behavioural therapy.

**References**


1. Mansfield KE, Sim J, Jordan JL, Jordan KP. A systematic review and meta-analysis of the prevalence of chronic widespread pain in the general population. Pain. 2016 Jan;157(1):55-64.
2. Jones GT, Atzeni F, Beasley M, Flüß E, Sarzi-Puttini P, Macfarlane GJ. The prevalence of fibromyalgia in the general population: a comparison of the American College of Rheumatology 1990, 2010, and modified 2010 classification criteria. Arthritis Rheumatol. 2015;67(2):568-75.
3. Burckhardt CS, Clark SR, Bennett RM. Fibromyalgia and quality of life: a comparative analysis. J Rheumatol. 1993;20(3):475-9.
4. Picavet HS, Hoeymans N. Health related quality of life in multiple musculoskeletal diseases: SF-36 and EQ-5D in the DMC3 study. Ann Rheum Dis. 2004;63(6):723-9.
5. Hughes G, Martinez C, Myon E, Taïeb C, Wessely S. The impact of a diagnosis of fibromyalgia on health care resource use by primary care patients in the UK: an observational study based on clinical practice. Arthritis Rheum. 2006;54(1):177-83.
6. Macfarlane GJ, Kronisch C, Dean LE, Atzeni F, Häuser W, Fluß E, Choy E, Kosek E, Amris K, Branco J, Dincer F, Leino-Arjas P, Longley K, McCarthy GM, Makri S, Perrot S, Sarzi-Puttini P, Taylor A, Jones GT. EULAR revised recommendations for the management of fibromyalgia. Ann Rheum Dis. 2017;76(2):318-328
7. von Korff M, Dunn KM. Chronic pain reconsidered. Pain. 2008;138(2):267-76.
8. Versus Arthritis Research Roadmap for pain: https://www.versusarthritis.org/media/1672/research-roadmap-pain.pdf
9. McBeth J, Prescott G, Scotland G, Lovell K, Keeley P, Hannaford P, McNamee P, Symmons DP, Woby S, Gkazinou C, Beasley M, Macfarlane GJ. Cognitive behaviour therapy, exercise, or both for treating chronic widespread pain. Arch Intern Med. 2012;172(1):48-57.
10. Beasley M, Prescott GJ, Scotland G, McBeth J, Lovell K, Keeley P, Hannaford PC, Symmons DP, MacDonald RI, Woby S, Macfarlane GJ. Patient-reported improvements in health are maintained 2 years after completing a short course of cognitive behaviour therapy, exercise or both treatments for chronic widespread pain: long-term results from the MUSICIAN randomised controlled trial. RMD Open. 2015;1(1):e000026.
11. Bernardy K, Klose P, Welsch P, Häuser W. Efficacy, acceptability and safety of cognitive behavioural therapies in fibromyalgia syndrome - A systematic review and meta-analysis of randomized controlled trials. Eur J Pain. 2018;22(2):242-260.

12. McBeth J, Macfarlane GJ, Benjamin S, Silman AJ. Features of somatization predict the onset of chronic widespread pain: results of a large population- based study. Arthritis Rheum. 2001;44(4):940-6.
13. Gupta A, Silman AJ, Ray D, Morriss R, Dickens C, MacFarlane GJ, Chiu YH, Nicholl B, McBeth J. The role of psychosocial factors in predicting the onset of chronic widespread pain: results from a prospective population-based study. Rheumatology (Oxford). 2007;46(4):666-71.
14. Macfarlane GJ, Beasley M, Prescott G, McNamee P, Keeley P, Artus M, McBeth J, Hannaford P, Jones GT, Basu N, Norrie J, Lovell K. The Maintaining Musculoskeletal Health (MAmMOTH) Study: Protocol for a randomised trial of cognitive behavioural therapy versus usual care for the prevention of chronic widespread pain. BMC Musculoskelet Disord. 2016 Apr 26;17:179.
15. Wolfe F, Smythe HA, Yunus MB, Bennett RM, Bombardier C, Goldenberg DL, Tugwell P, Campbell SM, Abeles M, Clark P, et al. The American College of Rheumatology 1990 Criteria for the Classification of Fibromyalgia. Report of the Multicenter Criteria Committee. Arthritis Rheum. 1990;33(2):160-72.
16. Kellner R. *Abridged Manual of the Illness Attitudes Scale*. University of New Mexico, 1983.
17. Othmer E, DeSouza C. A screening test for somatization disorder (hysteria). Am J Psychiatry. 1985;142(10):1146-9.
18. Jenkins CD, Stanton BA, Niemcryk SJ, Rose RM. A scale for the estimation of sleep problems in clinical research. J Clin Epidemiol. 1988;41(4):313-21.
19. Wolfe F, Clauw DJ, Fitzcharles MA, Goldenberg DL, Häuser W, Katz RS, Mease P, Russell AS, Russell IJ, Winfield JB. Fibromyalgia criteria and severity scales for clinical and epidemiological studies: a modification of the ACR Preliminary Diagnostic Criteria for Fibromyalgia. J Rheumatol. 2011;38(6):1113-22.
20. Goldberg DP, Williams P. A user's guide to the General Health Questionnaire. Windsor, ON, Canada: NFER-NELSON; 1988.
21. Chalder T, Berelowitz G, Pawlikowska T, Watts L, Wessely S, Wright D, Wallace EP. Development of a fatigue scale. J Psychosom Res. 1993;37(2):147-53.
22. Herdman M, Gudex C, Lloyd A, Janssen M, Kind P, Parkin D, Bonsel G, Badia X. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). Qual Life Res. 2011;20(10):1727-36.
23. Coast J, Peters TJ, Natarajan L, Sproston K, Flynn T. An assessment of the construct validity of the descriptive system for the ICECAP capability measure for older people. Qual Life Res. 2008;17(7):967-76.

24. Curtis L. Unit Costs of Health and Social Care (2014). Kent, UK: Personal Social Services Research Unit, University of Kent, 2014. https://www.gov.uk/government/publications/nhs-reference-costs-2013-to-2014
25. van Hout B, Janssen MF, Feng YS, 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.
26. 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.
27. Brand J, van Buuren S, le Cessie S, van den Hout W. Combining multiple imputation and bootstrap in the analysis of cost-effectiveness trial data. Stat Med. 2019;38(2):210-220.
28. Blackburn-Munro G. Hypothalamo-pituitary-adrenal axis dysfunction as a contributory factor to chronic pain and depression. Curr Pain Headache Rep. 2004;8 (2):116-24.
29. Jones GT, Power C, Macfarlane GJ. Adverse events in childhood and chronic widespread pain in adult life: Results from the 1958 British Birth Cohort Study. Pain. 2009;143(1-2):92-6.
30. Glette M, Stiles TC, Borchgrevink PC, Landmark T. The Natural Course of Chronic Pain in a General Population: Stability and Change in an Eight-Wave Longitudinal Study Over Four Years (the HUNT Pain Study). J Pain. 2019:S1526-5900(19)30845-4.
31. Landmark T, Romundstad P, Butler S, Kaasa S, Borchgrevink P. Development and course of chronic widespread pain: the role of time and pain characteristics (the HUNT pain study). Pain. 2019;160(9):1976-1981.
32. Zachariae R, Lyby MS, Ritterband LM, O'Toole MS. Efficacy of internet- delivered cognitive-behavioral therapy for insomnia - A systematic review and meta-analysis of randomized controlled trials. Sleep Med Rev. 2016 Dec;30:1-10.
33. Carlbring P, Andersson G, Cuijpers P, Riper H, Hedman-Lagerlöf E. Internet- based vs. face-to-face cognitive behavior therapy for psychiatric and somatic disorders: an updated systematic review and meta-analysis. Cogn Behav Ther. 2018;47(1):1-18.
34. Rutledge T, Atkinson JH, Holloway R, Chircop-Rollick T, D'Andrea J, Garfin SR, Patel S, Penzien DB, Wallace M, Weickgenant AL, Slater M. Randomized Controlled Trial of Nurse-Delivered Cognitive-Behavioral Therapy Versus Supportive Psychotherapy Telehealth Interventions for Chronic Back Pain. J Pain. 2018;19(9):1033-1039.

**Table 1: Baseline characteristics by treatment arm in the ITT population**

| **Characteristic** | **Randomised groups** | |
|---|---|---|
| | **tCBT (n=500)** | **Usual care (n=496)** |
| | *Median (IQR)* | *Median (IQR)* |
| Age (years) | 58.8 (47.7 – 68.7) | 59.5 (47.9 – 68.9) |
| | *N (%)* | *N (%)* |
| Gender | | |
| *Male* | 209 (41.8) | 204 (41.0) |
| *Female* | 291 (58.2) | 292 (58.9) |
| Employment status | | |
| *Working (full or part-time)* | 277 (55.4) | 244 (49.2) |
| *Unable to work because of health* | 18 (3.6) | 30 (6.0) |
| *Retired* | 168 (33.6) | 177 (35.7) |
| *Other* | 37 (7.4) | 45 (9.1) |
| CWP risk profile: | | |
| Illness behaviour score > 4 | | |
| *No* | 1 (0.2) | 2 (0.5) |
| *Yes* | 498 (99.6) | 494 (99.5) |
| *Not known** | 1 (0.2) | 0 (0.0) |
| Somatic Symptoms Scale score > 2 | | |
| *No* | 472 (94.4) | 462 (93.1) |
| *Yes* | 28 (5.6) | 34 (6.9) |
| Sleep problems score > 4 | | |
| *No* | 1 (0.2) | 2 (0.4) |
| *Yes* | 499 (99.8) | 493 (99.4) |
| *Not known** | 0 (0.0) | 1 (0.2) |
| CWP risk profile factors present (N) | | |
| *2* | 474 (94.8) | 466 (94.0) |
| *3* | 26 (5.2) | 30 (6.0) |
| | *Median (IQR) [n**]* | *Median (IQR) [n]* |
| Psychological distress (GHQ) | 1 (0 – 4) [499] | 1 (0 – 4) [494] |
| Quality of Life (EQ-5D-5L utility score) | 0.74 (0.65 – 0.80) [499] | 0.74 (0.64 – 0.80) [496] |
| ICECAP-A | 0.91 (0.81 – 0.95) [495] | 0.90 (0.79 – 0.95) [491] |
| Fibromyalgia Research Criteria | | |
| *WPI* | 3 (1 - 4) [499] | 2 (1 – 4) [492] |
| *SSS* | 4 (3 – 6) [497] | 4 (3 – 5) [494] |

EQ-5D-5L: Euroqol questionnaire – five dimensions – five levels; FRC: fibromyalgia research criteria; GHQ: general health questionnaire; ICECAP-A: ICEpop CAPability measure for Adults; IQR: interquartile range; SD: standard deviation; SSS: symptom severity scale; tCBT: telephone cognitive behavioural therapy; WPI: widespread pain index.

* Where individuals completed half or fewer items, the score was classified as not known, but individuals could still be eligible for recruitment based on their responses to other items answered.
** The number of persons for whom a scale score could be calculated

**Table 2: Outcomes by treatment arm at 12 months (Intention to treat analysis)**

| Characteristic | Randomised groups | |
|---|---|---|
| | **tCBT (n=500)** | **Usual care (n=496)** |
| | | |
| *Primary Outcome* | *N (%)* | *N (%)* |
| Chronic widespread pain | | |
| *No* | 315 (82.0) | 364 (82.5) |
| *Yes* | 69 (18.0) | 77 (17.5) |
| *Secondary Outcome* | | |
| Global impression of change | | |
| *Very much better* | 24 (6.5) | 15 (3.5) |
| *Much better* | 88 (23.7) | 59 (13.8) |
| *A little better* | 90 (24.3) | 84 (19.6) |
| *No change* | 83 (22.4) | 126 (29.4) |
| *A little worse* | 65 (17.5) | 119 (27.7) |
| *Much worse* | 18 (4.9) | 23 (5.4) |
| *Very much worse* | 3 (0.8) | 3 (0.7) |
| Pain Reported | | |
| *No* | 79 (20.6) | 68 (15.4) |
| *Yes* | 305 (79.4) | 373 (84.6) |
| CWP risk profile | | |
| Somatic symptoms score | | |
| *0* | 210 (56.5) | 228 (52.8) |
| *1* | 103 (27.7) | 123 (28.5) |
| *2-5* | 59 (15.9) | 81 (18.8) |
| Illness behaviour score *Mean (SD)[n*]* | 8.21 (4.04) [371] | 8.96 (4.19) [431] |
| Sleep problems score *Mean (SD)[n]* | 8.20 (4.89) [373] | 9.20 (5.16) [432] |
| Psychological Distress (GHQ score) | | |
| *0* | 201 (54.5) | 202 (46.8) |
| *1* | 59 (16.0) | 54 (12.5) |
| *2-5* | 68 (18.4) | 113 (26.2) |
| *6-12* | 41 (11.1) | 63 (14.6) |
| | *Mean (SD) [n]* | *Mean (SD) [n]* |
| Chalder Fatigue Score | 12.6 (4.5) [370] | 13.6 (4.4) [433] |
| | *Median (IQR) [n]* | *Median (IQR) [n]* |
| Quality of Life (EQ-5D utility score) | 0.74 (0.66 – 0.84) [371] | 0.74 (0.65 – 0.82) [435] |
| ICECAP-A | 0.91 (0.82 – 0.97) [368] | 0.89 (0.78 – 0.95) [429] |
| Fibromyalgia Research Criteria | | |
| *WPI* | 2 (1 - 4) [366] | 2 (1 – 4) [427] |
| *SSS* | 3 (2 – 5) [369] | 4 (2 – 5) [431] |

EQ-5D-5L: Euroqol questionnaire – five dimensions – five levels; FRC: fibromyalgia research criteria; GHQ: general health questionnaire; ICECAP-A: ICEpop CAPability measure for Adults; IQR: interquartile range; SD: standard deviation; SSS: symptom severity scale; tCBT: telephone cognitive behavioural therapy; WPI: widespread pain index. * The number of persons for whom a scale score could be calculated

**Table 3: Summary of the primary and secondary outcomes across follow-up points[1]**

| Outcome | | Time point (months) | Analysis method (effect size) | Adjusted* effect size (95% CI) | p-value | Unadjusted effect size (95% CI) | p-value |
|---|---|---|---|---|---|---|---|
| *Primary Outcome* | | | | | | | |
| CWP | | 3 | Logistic regression (OR) | 1.08 (0.74, 1.58) | 0.691 | 1.07 (0.75, 1.53) | 0.716 |
| CWP (per protocol) | | | | 1.15 (0.75, 1.75) | 0.519 | 1.18 (0.77, 1.66) | 0.522 |
| CWP (with multiple imputation) | | | | 1.06 (0.74, 1.54) | 0.749 | 1.05 (0.72, 1.53) | 0.816 |
| **CWP**** | | **12** | | **1.05 (0.75, 1.48)** | **0.771** | **1.04 (0.72, 1.48)** | **0.849** |
| CWP (per protocol) | | | | 1.11 (0.81, 1.50) | 0.519 | 1.09 (0.74, 1.60) | 0.673 |
| CWP (with multiple imputation) | | | | 1.04 (0.75, 1.45) | 0.982 | 1.03 (0.74, 1.42) | 0.964 |
| CWP | | 24 | | 0.85 (0.68, 1.07) | 0.163 | 0.84 (0.61, 1.18) | 0.317 |
| CWP (per protocol) | | | | 0.85 (0.64, 1.12) | 0.241 | 0.84 (0.58, 1.20) | 0.330 |
| CWP (with multiple imputation) | | | | 0.85 (0.66, 1.09) | 0.220 | 0.84 (0.65, 1.09) | 0.196 |
| CWP | | 3,12,24 | GEE (OR) | 1.00 (0.96, 1.04) | 0.923 | 1.00 (0.96, 1.04) | 0.835 |
| *Secondary Outcomes* | | | | | | | |
| Global impression of change[2] | | 3 | Ordinal Logistic Regression (OR) | 0.42 (0.32, 0.55) | <0.001 | 0.43 (0.34, 0.56) | <0.001 |
| | | 12 | | 0.51 (0.39, 0.67) | <0.001 | 0.53 (0.41, 0.68) | <0.001 |
| | | 24 | | 0.55 (0.43, 0.70) | <0.001 | 0.58 (0.45, 0.73) | <0.001 |
| CWP risk profile | Somatic symptom score | 3 | Ordinal logistic regression (OR) | 0.79 (0.60, 1.03) | 0.084 | 0.83 (0.64, 1.08) | 0.173 |
| | | 12 | | 0.86 (0.71, 1.04) | 0.112 | 0.85 (0.65, 1.11) | 0.237 |
| | | 24 | | 0.81 (0.59, 1.12) | 0.206 | 0.90 (0.67, 1.21) | 0.498 |
| | Illness behaviour score | 3 | Linear regression (Mean difference) | -0.17 (-0.58, 0.24) | 0.385 | -0.25 (-0.79, 0.29) | 0.360 |
| | | 12 | | -0.81 (-1.54, -0.09) | 0.030 | -0.74 (-1.32, -0.17) | 0.011 |
| | | 24 | | -1.25 (-2.15, -0.35) | 0.010 | -1.20 (-1.83, -0.58) | <0.001 |
| | Sleep problems score | 3 | Linear regression (Mean difference) | -0.62 (-1.26, 0.02) | 0.057 | -0.62 (-1.31, 0.08) | 0.081 |
| | | 12 | | -0.95 (-1.48, -0.42) | 0.002 | -1.00 (-1.70, -0.30) | 0.005 |
| | | 24 | | -0.51 (-1.25, 0.23) | 0.161 | -0.52 (-1.39, 0.16) | 0.117 |

[1] Analyses shaded in grey favour tCBT over usual care at pre-specified significance level for secondary outcomes (p<0.01). Except for EQ-5D-5L, mean differences less than zero and odds ratios less than one favour the treatment group.

[2] OR of one point increase in global impression of change score (worsening of health)

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Psychological distress (GHQ) | | 3 | Ordinal logistic regression (OR) | 0.55 (0.43, 0.69) | <0.001 | 0.58 (0.45, 0.76) | <0.001 |
| | | 12 | | 0.65 (0.50, 0.86) | 0.002 | 0.70 (0.54, 0.90) | 0.007 |
| | | 24 | | 0.76 (0.60, 0.96) | 0.024 | 0.74 (0.56, 0.98) | 0.037 |
| Chalder fatigue score | | 3 | Linear Regression (Mean difference) | -1.36 (-2.10, -0.64) | 0.001 | -1.40 (-1.97, -0.82) | <0.001 |
| | | 12 | | -1.02 (-1.63, -0.42) | 0.003 | -1.03 (-1.64, -0.42) | 0.001 |
| | | 24 | | -0.93 (-1.62, -0.23) | 0.012 | -0.93 (-1.58, -0.27) | 0.006 |
| Quality of Life (EQ-5D-5L utility score) | | 3 | Linear Regression (Mean difference) | 0.009 (-0.009, 0.028) | 0.304 | 0.021 (-0.004, 0.046) | 0.101 |
| | | 12 | | 0.024 (0.009, 0.040) | 0.004 | 0.037 (0.010, 0.064) | 0.007 |
| | | 24 | | 0.030 (0.009, 0.050) | 0.008 | 0.040 (0.011, 0.069) | 0.007 |
| ICECAP-A tariff | | 3 | Ordinal logistic regression (OR) | 1.14 (0.89, 1.48) | 0.304 | 1.17 (0.86, 1.59) | 0.323 |
| | | 12 | | 1.39 (0.94, 2.04) | 0.096 | 1.39 (1.01, 1.91) | 0.042 |
| | | 24 | | 0.88 (0.67, 1.15) | 0.338 | 0.99 (0.70, 1.41) | 0.966 |
| Fibromyalgia criteria | Widespread Pain Index | 3 | Poisson Regression (IRR) | 0.98 (0.90, 1.07) | 0.698 | 1.01 (0.93, 1.10) | 0.771 |
| | | 12 | | 0.88 (0.80, 0.98) | 0.018 | 0.92 (0.84, 0.99) | 0.036 |
| | | 24 | | 0.88 (0.78, 0.98) | 0.022 | 0.92 (0.84, 1.00) | 0.058 |
| | Symptom Severity Scale | 3 | Linear Regression (Mean difference) | -0.28 (-0.52, -0.04) | 0.026 | -0.25 (-0.57, 0.65) | 0.118 |
| | | 12 | | -0.52 (-0.75, -0.28) | <0.001 | -0.59 (-0.91, -0.27) | <0.001 |
| | | 24 | | -0.29 (-0.55, -0.02) | 0.040 | -0.28 (-0.61, 0.05) | 0.100 |

CI: confidence interval; CWP: chronic widespread pain; EQ5D-5D-5L: Euroqol questionnaire – five dimensions – five levels; GEE: generalised estimating equations with unstructured correlation structure; GHQ: general health questionnaire; ICECAP-A: ICEpop CAPability measure for Adults; IRR: incident rate ratio; OR: odds ratio;

*Adjusted analyses control for the number of risk factors (two or three), age, gender, baseline score (if applicable) and centre (random effect). Analyses are intention-to-treat unless otherwise stated. ** Primary outcome.

**Table 4: Adjusted[3] mean incremental costs, incremental QALYs, and incremental cost-effectiveness ratio over 24 months between tCBT vs usual care**

| Analysis | Mean costs, (95% CI) | | Mean QALYs, (95% CI) | | Incremental mean costs, £ (95% CI)[4] | Incremental mean QALYs (95% CI) | ICER (£/QALY) |
|---|---|---|---|---|---|---|---|
| | tCBT | Usual care | tCBT | Usual care | | | |
| Imputed dataset/ITT analysis (NHS perspective)[5] | 3094.68 (1775.65-9074.15) | 3052.38 (1735.77-8567.24) | 1.254 (1.238-1.270) | 1.231 (1.215-1.245) | 42.30 (-451.19-597.90) | 0.023 (0.007-0.039) | 1,828 |
| SA: Complete cases (NHS perspective)[6] | 2684.53 (1817.69-5221.86) | 2454.67 (1645.66-4769.87) | 1.444 (1.415-1.471) | 1.420 (1.392-1.447) | 229.86 (-228.74-734.09) | 0.024 (-0.005-0.053) | 9,608 |
| SA: Imputed dataset (NHS + private care perspective) | 4239.22 (2135.82-15332.80) | 4149.10 (2110.98-14039.06) | 1.253 (1.238-1.270) | 1.231 (1.215-1.247) | 90.12 (-475.79-772.98) | 0.022 (0.007-0.039) | 4,022 |
| SA: Imputed dataset using actual trial expenses (NHS perspective)[7] | 3128.61 (1809.54-9164.04) | 3027.54 (1734.31-8587.83) | 1.254 (1.238-1.270) | 1.231 (1.215-1.245) | 101.07 (-373.14-641.98) | 0.023 (0.007-0.039) | 4,367 |
| SA: Imputed dataset using the cost of a complete tCBT course (NHS perspective)[8] | 3314.57 (1966.93-9059.99) | 2960.98 (1729.67-7781.19) | 1.254 (1.238-1.270) | 1.231 (1.215-1.245) | 353.59 (-80.46-1,238.07) | 0.023 (0.007-0.039) | 15,280 |
| SA: Imputed dataset using ICECAP (NHS perspective)[9] | 4659.66 (1764.56-10400.07) | 4787.56 (1815.05-10632.20) | 1.288 (1.278-1.297) | 1.275 (1.266-1.284) | -127.90 (-603.19-545.33) | 0.013 (0.003-0.023)[10] | NA |

[3] Adjusted for baseline differences (age, gender, number of risk factors present, employment status, centre, baseline EQ-5D health utility score and baseline cost).

[4] Bootstrapped non-parametric 95% confidence interval ($2.5^{th}$/$97.5^{th}$ centile). Generalised linear model with γ distribution and log-link function to estimate incremental costs and generalised linear model with Poisson distribution and power 0.5 link function to estimate incremental QALYs/years of full capacity. Discounted at 3.5% per year

[5] Imputed dataset is the ITT analysis. Missing values were imputed to account for all participants included in the ITT analysis.

[6] 593 complete cases were included (tCBT, n=297 and usual care, n=326). Complete cases are those with no missing data on cost and health utility at each time point.

[7] Included the actual trial expenses per tCBT participant, £301. This was estimated using the lump-sum trial expenses incurred by therapists, including therapists' training and tCBT delivery.

[8] Included the cost of a complete tCBT course per participant, £443. Time spent by therapist, training and supervision were included. The total time spent by the therapist was estimated by assuming that all tCBT participants attended a complete tCBT course consisting of 9 sessions.

[9] Adjusted for baseline differences (age, gender, number of risk factors present, employment status, centre, baseline ICECAP value and baseline cost).

[10] Incremental years of full capability.

QALYs, quality-adjusted life-years; ICER, incremental cost-effectiveness ratio; NA, not applicable; SA, sensitivity analysis.

**Supplementary Text**

Methods

The Scottish Index of multiple deprivation was used to determine the quintile of deprivation of participants based on their area of residence, assessed by their postcode (SIMD, 2016).

*Details of Secondary Outcomes*

Secondary outcomes collected in the study were as follows:

- Global Impression of Change (of health since entering the trial), a single-item measure of seven six categories (very much better, much better, a little better, no change, a little worse, much worse and very much worse), although the latter two categories were combined for analysis due to low numbers;
- the presence of pain over the last month and number of pain sites (measured by the Widespread Pain Index (score 0-19) of the "research" (or 2010 revised) criteria for fibromyalgia (20). In addition, we measured the Symptom Severity Scale (SSS: score 0-12) which are also part of the fibromyalgia criteria set. Participants with WPI ≥7 & SSS ≥ 5, or WPI 3-6 & SSS ≥9, and who reported having such symptoms for at least 3 months, meet criteria for fibromyalgia.
- the "risk profile" for CWP as assessed by the Illness Behaviour Subscale of the Illness Attitudes Scale (16), the Somatic Symptom Scale score (but excluding items on pain) (17), and the Sleep Problem Scale (18). The Illness Attitudes Scale measures attitudes and concerns about illness and health. A study using principal component analysis (Speckens et al, 1996) showed that the IAS consisted of two subscales, one of which related to illness behaviour (6 items), scoring from 0-24, with higher scores associated with undertaking specific behaviours. The Somatic Symptom Scale was originally devised as a screening tool for somatisation and consists of 5 non-pain items (0-5, with higher scores indicating more somatic symptoms). The Sleep Problem Scale consists of four items measuring sleep problems over the past four weeks with score range 0-20, higher scores indicating greater frequency of sleep problems;
- psychological distress measured using the 12-item General Health Questionnaire (GHQ)(20) and analysed using an ordinal model with the categories 0 (least distress), 1, 2-5 and 6-12 (most distress);
- fatigue measured using the Chalder Fatigue Scale (11 items with scores 0-33, higher scores representing more disabling and severe fatigue) (21);
- quality of life measured using the five-item, five level EQ-5D-5L (-0.59 representing the worst possible quality of life and 1 the best possible) (22);
- capability using the 5-item ICECAP-A (ICEpop CAPability measure for Adults) which focusses on wellbeing, and analysed using an ordinal model with categories 0-0.49 (worst quality of life), 0.5-0.79 and 0.8-1.0 (best quality of life) (23);

*Sensitivity analyses*

Sensitivity analyses were conducted for the primary outcome only. These included an analysis excluding participants who did not complete the active intervention (per protocol analysis) and an analysis using multiple imputation (see Royston, 2004). For the per protocol analysis, participants in the intervention group were included if they had the initial assessment with the therapist and it was mutually agreed to stop the treatment, or if they had the initial assessment plus at least 2 more sessions with the therapist. Missing values for CWP at each time point were imputed using the *mi* package in STATA using the following variables: age, gender, number of risk factors and GP practice. Twenty imputed datasets were created, using an adjusted logistic regression model. An additional analysis for CWP incorporating all three follow-up time points in one model was also conducted using generalised estimating equations using an unstructured correlation structure (see Zeger et al, 1988). The model was adjusted for covariates and results expressed as an OR with 95% CI.

Additional requirements in CONSORT reporting of trials

Randomisation: The randomisation was undertaken by a member of the study team contacting, using internet or telephone, the trial randomisation centre at the Centre for Healthcare Randomised Trials at the University of Aberdeen. The participants were informed of the allocated group during the consent/randomisation phone call. The study statistician was blinded to which group received the “active” treatment until the statistical analysis had been completed.

Generalisability: The trial recruited from very different areas of Scotland. It included urban areas in Glasgow and Aberdeen with very different levels of deprivation, rural areas in Aberdeenshire and remote areas across the Highlands. Recruitment through a population-sampling frame maximises generalisability. Our previous (qualitative) work in terms of telephone delivery of CBT has shown that this can improve access both in remote and rural areas (because care can be obtained without long distance travel) and in urban areas (since it overcomes, for example, difficulties in getting time off work or in arranging suitable care for dependents) (Bee et al, 2010; Bee et al, 2016).

Harms: It was not envisaged that the intervention would lead to harms, but procedures were designed to support any participants who became distressed during the sessions.

Supplementary Text References

Bee PE, Lovell K, Lidbetter N, Easton K, Gask L. You can't get anything perfect: "user perspectives on the delivery of cognitive behavioural therapy by telephone". Soc Sci Med. 2010;71(7):1308-1315.

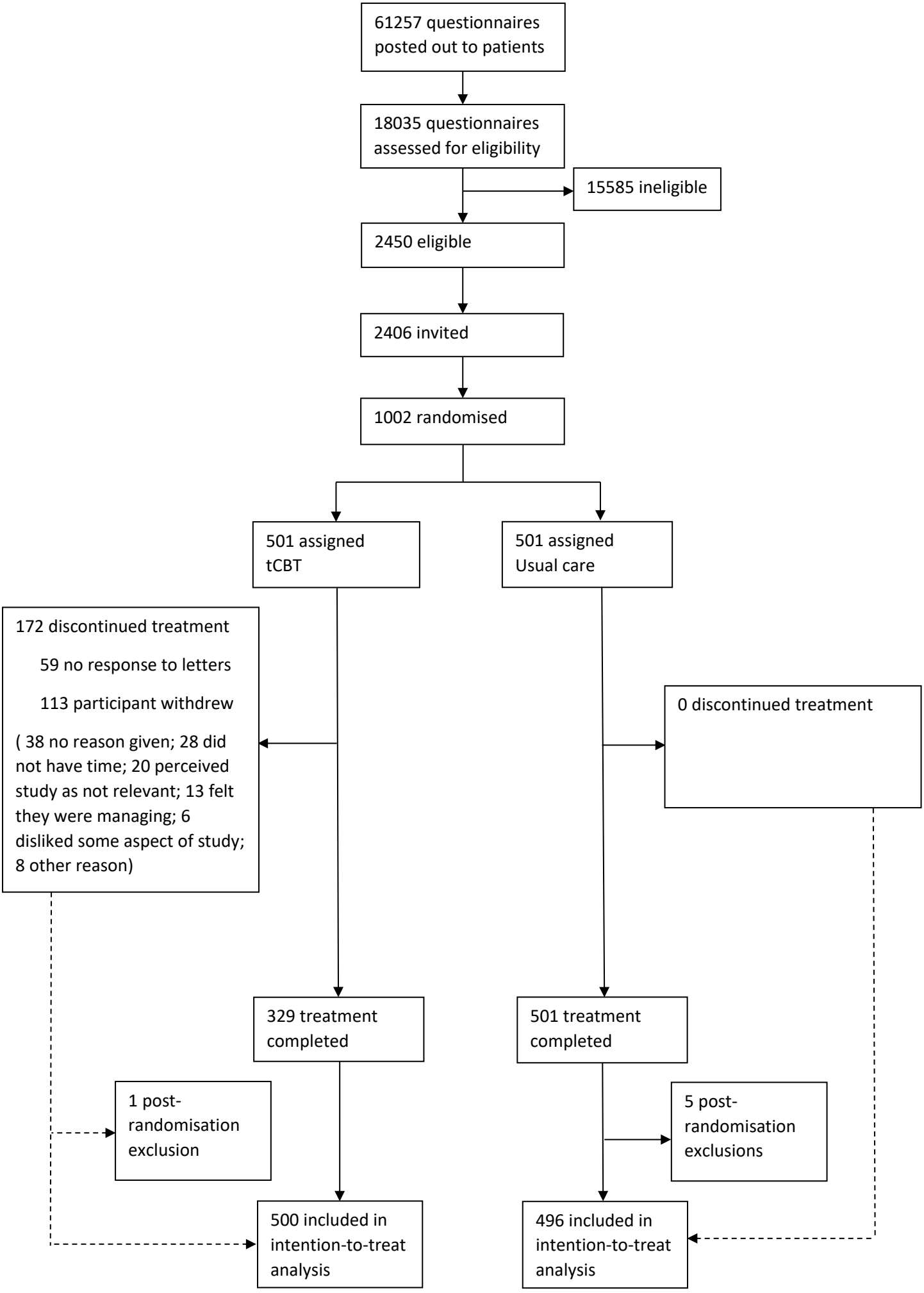
Bee P, McBeth J, Macfarlane GJ, Lovell K. Managing chronic widespread pain in primary care: a qualitative study of patient perspectives and implications for treatment delivery. BMC Musculoskelet Disord. 2016;17(1):354.

Royston P. Multiple imputation of missing values. Stata Journal 2004 4(3): 227-241.

Scottish Index of Multiple Deprivation (SIMD) 2016: https://www2.gov.scot/SIMD

Speckens AE, Spinhoven P, Sloekers PPA, Bolk JH, van Hemert AM. A validation study of the Whitley Index, the Illness Attitude Scales, and the Somatosensory Amplification Scale in general medical and general practice patients. J Psychosom Res 1996;40:95-104

Zeger SL, Liang KY, Albert PS. Models for longitudinal data: a generalized estimating equation approach. Biometrics. 1988;44(4):1049-60.


61257 questionnaires posted out to patients
18035 questionnaires assessed for eligibility
15585 ineligible
2450 eligible
2406 invited
1002 randomised
501 assigned tCBT
501 assigned Usual care
172 discontinued treatment
59 no response to letters
113 participant withdrew
( 38 no reason given; 28 did not have time; 20 perceived study as not relevant; 13 felt they were managing; 6 disliked some aspect of study; 8 other reason)
0 discontinued treatment
329 treatment completed
501 treatment completed
1 post-randomisation exclusion
5 post-randomisation exclusions
500 included in intention-to-treat analysis
496 included in intention-to-treat analysis