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.