Effectiveness of a symptom-clinic intervention delivered by general practitioners with an extended role for people with multiple and persistent physical symptoms in England: the Multiple Symptoms Study 3 pragmatic, multicentre, parallel-group, individually randomised controlled trial

Christopher Burton, Cara Mooney, Laura Sutton, David White, Jeremy Dawson, Aileen R Neilson, Gillian Rowlands, Steve Thomas, Michelle Horspool, Kate Fryer, Monica Greco, Tom Sanders, Ruth E Thomas, Cindy Cooper, Emily Turton, Waquas Waheed, Jonathan Woodward, Ellen Mallender, Vincent Deary

Summary

Background People with multiple and persistent physical symptoms have impaired quality of life and poor experiences of health care. We aimed to evaluate the effectiveness of a community-based symptom-clinic intervention in people with multiple and persistent physical symptoms, hypothesising that this symptoms clinic plus usual care would be superior to usual care only.

Methods The Multiple Symptoms Study 3 was a pragmatic, multicentre, parallel-group, individually randomised controlled trial conducted in 108 general practices in the UK National Health Service in four regions of England between Dec 6, 2018, and June 30, 2023. Participants were individually randomised (1:1) to the symptom-clinic intervention plus usual care or to usual care only via a computer-generated, pseudo-random list stratified by trial centre. Allocation was done by the trial statistician and concealed with a centralised, web-based randomisation system; masking participants was not possible due to the nature of the intervention. The symptom-clinic intervention was a sequence of up to four medical consultations that aimed to elicit a detailed clinical history, fully hear and validate the participant, offer rational explanations for symptoms, and assist the participant to develop ways of managing their symptoms; it was delivered by general practitioners with an extended role. The primary outcome was Patient Health Questionnaire-15 (PHQ-15) score 52 weeks after randomisation, analysed by intention to treat. The trial is registered on the ISRCTN registry (ISRCTN57050216).

Findings 354 participants were randomly assigned; 178 (50%) were assigned to receive the community-based symptoms clinic plus usual care and 176 (50%) were assigned to receive usual care only. At the primary-outcome point of 52 weeks, PHQ-15 scores were 14·1 (SD 3·7) in the group receiving usual care and 12·2 (4·5) in the group receiving the intervention. The adjusted between-group difference of –1·82 (95% CI –2·67 to –0·97) was statistically significantly in favour of the intervention group (p<0·0001). There were 39 adverse events in the group receiving usual care and 36 adverse events in the group receiving the intervention. There were no statistically significant between-group differences in the proportion of participants who had non-serious adverse events (–0·03, 95% CI –0·11 to 0·05) or serious adverse events (0·02, –0·02 to 0·07). No serious adverse event was deemed to be related to the trial intervention.

Interpretation Our symptom-clinic intervention, which focused on explaining persistent symptoms to participants in order to support self-management, led to sustained improvement in multiple and persistent physical symptoms.

Funding UK National Institute for Health and Care Research.

Lancet 2024; 403: 2619–29

See **Comment** page 2568

Division of Population Health (Prof C Burton MD, K Fryer PhD), **Clinical Trials Research Unit** (C Mooney MSc, D White MPH, Prof C Cooper PhD, E Turton MA, J Woodward MSc), **Sheffield Centre for Health and Related Research** (L Sutton PhD, Prof J Dawson PhD), **and Sheffield University Management School** (Prof J Dawson), **University of Sheffield, Sheffield, UK; Edinburgh Clinical Trials Unit, Usher Institute, University of Edinburgh, Edinburgh, UK** (A R Neilson MSc); **Public Health Sciences Institute, Newcastle University, Newcastle upon Tyne, UK** (G Rowlands MD); **South Yorkshire Integrated Care Board, Sheffield, UK** (S Thomas MBBS, M Horspool PhD); **Department of Social and Policy Sciences, University of Bath, Bath,** UK (Prof M Greco PhD); **Department of Social Work, Education and Community Wellbeing** (Prof T Sanders PhD) **and Department of Psychology** (Prof V Deary PhD), **Northumbria University, Newcastle upon Tyne, UK; Centre for Healthcare Trials, Health Services Research Unit, University of Aberdeen, Aberdeen, UK** (R E Thomas PhD); **Centre for Primary Care, Institute of Population Health, University of Manchester, Manchester, UK** (W Waheed PhD); **Chesterfield, UK** (E Mallender BA)

Correspondence to:
Prof Christopher Burton, Division of Population Health, University of Sheffield, Sheffield S1 4DA, UK
chris.burton@sheffield.ac.uk

Introduction

Persistent physical symptoms are common in all clinical settings. When disproportionate to detectable physical disease or representing so-called functional disorders, they constitute approximately a third of referrals from general practitioners (GPs) to specialists in the UK.[1] Furthermore, approximately 2% of adults in the UK have been shown to have multiple physical symptoms that substantially affect their quality of life.[2] Persistent physical symptoms have substantial costs for both health services and society.[3] Although the prognosis for isolated symptoms is relatively good, the presence of multiple symptoms, regardless of diagnosis, is a strong predictor of poor health status and high health-care use.[4] Classification and nomenclature for persistent physical symptoms is currently evolving, so we wanted to assess

Research in context

Evidence before this study

Persistent physical symptoms, which are disproportionate to any underlying disease, are common in all areas of medicine, including primary care. Developments in knowledge about persistent symptoms during the past 10 years suggest they should no longer be considered medically unexplained or simply be re-attributed to mental distress. Whether enhanced communication by general practitioners (GPs) can influence persistent physical symptoms is unclear. We used the Ovid interface to search PubMed from Jan 1, 2010, to Dec 31, 2023, for clinical trials of primary-care interventions for persistent physical symptoms published in English only. We used the multipurpose (.mp) qualifier with the search terms ("persistent adj2 symptoms" OR "medically unexplained" OR "functional adj2 (disorder$ or symptom$ or syndrome$)" OR "somatis$ or somatiz$ or somatof$") AND ("Primary care" OR "General Practi$" OR "Family adj (medicihysicianysicia$ or practi$)") AND ("Randomi$" OR "trial" OR "systematic review"). The search identified 243 papers, including three relevant systematic reviews and seven full, primary care-based trials, none of which evaluated GPs with an extended role.

Added value of this study

Our pragmatic, multicentre, parallel-group, individually randomised controlled trial evaluated an enhanced clinical-communication model for people with multiple and persistent physical symptoms delivered by GPs with an extended role in four English regions. The intervention was delivered as intended during a 3-month period; 12 months after enrolment, there was a statistically significant between-group difference in physical symptoms favouring the intervention. The number needed to treat for an improvement of twice the clinically important difference was five. There was no evidence of harm due to missed or delayed diagnosis due to the intervention. To our knowledge, this trial is the first to evaluate an intervention with GPs with an extended role for patients with multiple and persistent physical symptoms.

Implications of all the available evidence

Our results add to systematic-review findings that suggest possible benefit from intensive and enhanced consultation interventions. The evolution of clinical roles in primary care indicates that this approach of enhanced consultations delivered by GPs with an extended role is an opportunity to improve outcomes for people with persistent physical symptoms, a common and important condition.

both individual symptoms and symptom-based syndromes or functional disorders.[5,6] Although the term medically unexplained symptoms is still widely used, we share the opinion that it is no longer appropriate.[7]

People with persistent physical symptoms commonly have poor experiences of health care[8,9] and are often told that their symptoms cannot be explained.[10] However, developments during the past 10 years in interoception,[11] symptom perception,[12] and biopsychosocial integration[13] mean that persistent symptoms can be increasingly understood as entities in their own right[5] and usefully explained.[14] Meanwhile, the focus of medical investigation on finding pathology and attributing symptoms to that pathology means that reassurance from diagnostic tests is typically transient[15] and people with persistent physical symptoms often feel stuck, disbelieved, and helpless.[8]

We developed an extended consultation intervention for people with multiple and persistent physical symptoms that were disproportionate to diagnosed physical disease, which was delivered by GPs with an extended role—defined by the Royal College of General Practitioners as GPs who are "undertaking an activity that is beyond the scope of general practice and requires further training"—who typically see patients referred from multiple general practices.[16,17] The aim of this symptom-clinic intervention was to recognise and validate the experiences of a person, explain symptoms through co-production on the basis of research on symptoms published in the past 10 years,[14,18] and agree actions on the basis of the explanation to manage symptoms or limit their effect.[19] In this trial, we aimed to evaluate the effectiveness of this symptom-clinic intervention in people with multiple and persistent physical symptoms. We hypothesised that a community-based symptoms clinic plus usual care would be superior to usual care only.

Methods

Study design

The Multiple Symptoms Study 3 (MSS3) was a pragmatic, multicentre, parallel-group, individually randomised controlled trial. The trial was initially conducted in person for enrolment and delivery of the intervention but, after a pause between March 14, 2020, and Aug 24, 2020, due to the COVID-19 pandemic, it changed to an online video-call or telephone-call format for both enrolment and delivery of the intervention. MSS3 was conducted in 108 general practices in the UK National Health Service—initially in the three English regions of Yorkshire and the Humber, Greater Manchester, and Newcastle and Gateshead—between Dec 6, 2018, and June 30, 2023. A fourth participant-identification region was established in northwest London when the trial was being delivered remotely.

National Health Service research ethics approval was received from the Greater Manchester Central Research

Ethics Committee (reference 18/NW/0422) on June 25, 2018. There was an independent trial steering committee and an independent data monitoring and ethics committee.

People with lived experience of persistent physical symptoms were involved in the design and development of MSS3. Participant input was incorporated in the delivery of the project through representation in the internal trial-management group and as independent members of the trial steering committee. Participants in both groups received a £10 voucher on completion of the final outcome questionnaire.

Participants

Eligible participants were adults aged 18–69 years with multiple and persistent physical symptoms as evidenced by a Patient Health Questionnaire-15 (PHQ-15)[20] score between 10 and 20. This score represents moderate or moderately severe symptom burden. We excluded people with PHQ-15 score more than this range, as pilot work indicated that they were likely to be too severely affected by their symptoms to benefit from this moderate-intensity intervention. To be eligible, individuals were required to have at least one code for a syndrome defined by physical symptoms, such as irritable bowel syndrome or fibromyalgia, in their GP electronic health record and to have been referred to specialists at least twice in the preceding 36 months (later extended to 42 months due to pandemic lockdowns). Exclusion criteria were evidence in medical records of previous or current major illness likely to cause symptoms, inability to manage personal care or leave the home independently, substantial thoughts of self-harm (ie, a score of 3 on the self-harm question of the Patient Health Questionnaire-9 [PHQ-9]; appendix pp 73–74), difficulty conducting a health-care consultation in English without either a professional or family interpreter or other assistance, current pregnancy or pregnancy within the past 6 months, and currently undergoing specialist rehabilitation or psychological therapy.[21]

Participants were recruited in four stages. First, GP practices did a structured search on their electronic health records to identify potential participants on the basis of diagnostic and referral codes. The results of these searches were checked by a GP within the practice to avoid sending invitations to people for whom it would be inappropriate, including those for whom the GP believed symptoms were more likely to be attributable to an underlying medical condition. Second, practices sent a trial-participation invitation and information pack to people identified via this search. This pack contained the participant information sheet, a PHQ-15 questionnaire, and a reply form to be sent to the trial team. Third, individuals who expressed an interest in the trial and whose PHQ-15 score was in the eligible range received a screening call from a research assistant from the Clinical Trials Research Unit at the University of Sheffield to check for exclusion criteria and for availability to attend an enrolment session (before March, 2020) or take part in online enrolment (after September, 2020). An enrolment appointment was conducted either in person or online; in-person enrolment was conducted by a research nurse and online enrolment was conducted by a research asssistant from the Clinical Trials Research Unit. At this point participants gave written informed consent, baseline measures were completed, and randomisation was done. Gender was self-reported by participants as woman or man.

Randomisation and masking

At conclusion of enrolment, participants were individually randomly assigned (1:1) to the symptom-clinic intervention plus usual care or to usual care only via a computer-generated, pseudo-random list stratified by trial centre with random permuted blocks of varying sizes. Randomisation lists were generated by a statistician from the Clinical Trials Research Unit at the University of Sheffield who was not subsequently involved in the analysis. Allocation was made and concealed by a centralised, web-based randomisation system. Participants were informed of their allocation at their enrolment appointment; if assigned to the intervention, the first appointment at the symptom clinic was scheduled. Due to the nature of the intervention, masking participants to whether they were in the intervention group or in the usual care only group was not possible. The processes and timing of outcome collection were identical across both groups and were done by staff at the Clinical Trials Research Unit (including JW), who were masked to allocated group. Statistical analysis was done masked to treatment allocation.

See Online for appendix

Procedures

The symptom-clinic intervention was a sequence of up to four medical consultations that aimed to elicit a detailed clinical history, ensure that the participant's experience was fully heard and validated, offer rational explanations for symptoms, and assist the participant to develop ways of managing their symptoms. The treatment model can be summarised with the terms recognition, explanation, action, and learning (REAL; figure 1).[19] The symptom-clinic intervention was delivered by GPs with an extended role, whose work typically uses the skills of holism and of managing complexity and uncertainty, which are central to generalism. The extended role is often done by practitioners in "a setting outside their usual general practice and involves receiving referrals for assessment and treatment from outside their immediate practice".[17] GPs with an extended role were recruited by application and interview. In-person consultations were with participants from their own region, video consultations could be with participants from any region.

The symptom clinic consisted of up to four one-to-one consultations with a GP with an extended role during

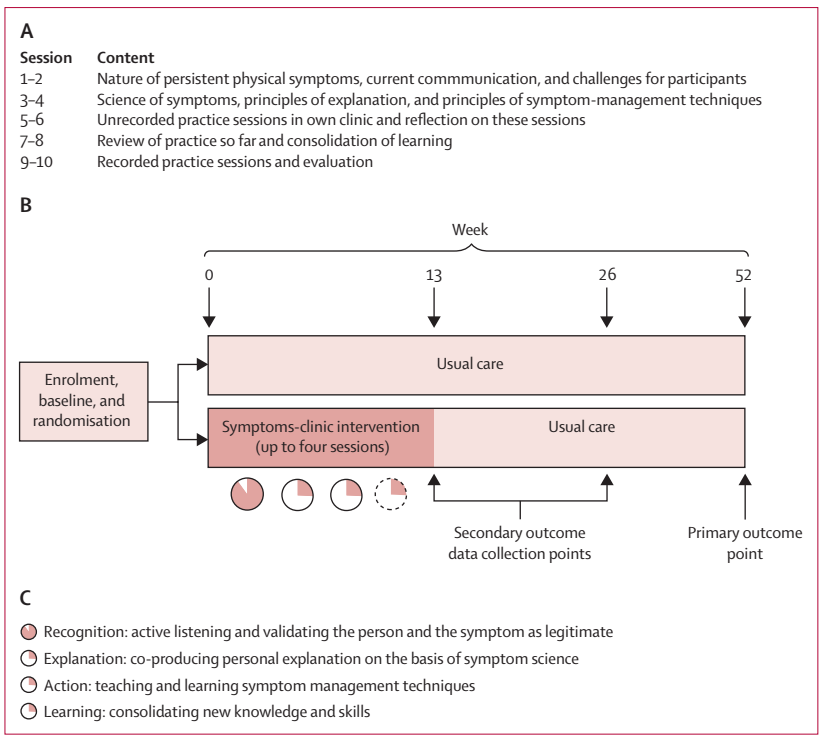

***Figure 1*: Summary of the Multiple Symptoms Study 3**
(A) Training of general practitioners with an extended role. (B) Trial design. (C) Symptoms-clinic intervention components.

approximately 8 weeks, varying from 6 weeks to 12 weeks. The initial consultation was approximately 50 min and the subsequent up to three medium consultations were approximately 15–20 min. Consultations typically occurred with intervals of 2 or more weeks between. GPs with an extended role had flexibility to increase the length of time between consultations if required. Participants attending the symptom-clinic intervention received an SMS reminder the day before each consultation and were offered a further appointment if they missed one. Participants who missed this further appointment were not offered any more appointments.

Protocol compliance in the intervention group was predefined (appendix p 35) as participants attending the initial consultation plus at least one follow-up session. We considered number of consultations to be the most robust measure of engagement with the intervention rather than duration of treatment in weeks, which could be affected by random events (eg, holidays or unavailability). We made a post-hoc adjudication as to whether discontinuation before the fourth session was planned or not by using completion of the second PSYCHLOPS measure, as this measure was part of the final consultation.

Consultations were structured to elicit a detailed medical history. During the first consultation, the participant and GP completed the psychological outcome profiles (PSYCHLOPS) measure[22] to focus on two specific symptoms. Physical examination was not routinely done, although GPs could conduct a targeted clinical examination if necessary. Psychosocial issues were discussed when participants mentioned them, with a focus on viewing these experiences in the context of persistent symptoms rather than attributing symptoms directly to them. GPs with an extended role were taught to be alert for symptoms suggesting undiagnosed pathology and were able to refer participants back to their usual GP for consideration of tests or treatment. During the trial, GPs with an extended role received supervision from CB or VD approximately every 2 months.

Throughout the study, participants in both groups continued to receive usual care from their regular GP practice. Participants had no contact with the GPs with an extended role who were involved in the trial other than during the intervention, except for receiving a copy of a letter sent by the GP with an extended role to their regular GP after the first and last consultations that described the formulation and any planned actions.

We audio-recorded all symptom-clinic consultations and used a stratified sample of 30%, representing consultations from each of the GPs with an extended role, for quality assurance and process evaluation. We assessed fidelity to the intervention using a set of 16 criteria developed from the REAL model[19] and evidencing them from the recordings using a traffic-light system in which green indicated that a criterion was clearly met, amber indicated that a criterion was possibly met, and red indicated that a criterion was not met. The process evaluation also included semi-structured interviews of 19 participants who received the symptom-clinic intervention. Potential interviewees were identified by CM, a trial manager who was not masked to allocation and was not involved in outcome-data collection or processing. Purposive sampling ensured a mix of participants who were currently working or not working and those with different predominant symptoms (eg, pain or fatigue). Potential interviewees were approached by KF, a qualitative researcher, via their preferred method of contact. Interviews were done by KF, who also assessed treatment fidelity. Details of the conduct and findings of the interviews have been published elsewhere.[19]

We used participant-reported outcome measures completed at 13 weeks, 26 weeks, and 52 weeks after randomisation via posted paper forms. If participants did not return these measures, they were followed up on three occasions via telephone call, SMS message, or posted reminder pack, depending on what contact details they provided and how they had responded previously. If paper forms were not returned, participants were offered the opportunity to complete these measures via telephone. We minimised risk of bias due to unmasking of allocation during telephone calls by stating at the beginning of the call that the individual should not reveal their allocation. We kept record of when unmasking happened and arranged for another member of the trial

team to complete subsequent telephone assessments for that participant.

The PHQ-15 consists of 15 symptoms for which participants were asked to report severity during the past 4 weeks as either 0 (not bothered at all), 1 (bothered a little), or 2 (bothered a lot). Two items in the PHQ-15 relate to symptoms that were not relevant to all participants (ie, menstrual symptoms and symptoms related to sexual intercourse); if these items were left blank, the total score was adjusted on the basis of the mean of remaining entries. The PHQ-15 has been reported as having an MCID of 2·3 points,[23] although this finding was not published at the time our trial was designed. Data for age, ethnicity, first language, education, 5-Level EuroQol-5 Dimension (EQ-5D-5L) score,[24] European Health Literacy Survey-6 (HLS-EU6) score,[25] ICEpop Capability Measure for Adults (ICECAP-A) score,[26] Patient Global Impression of Change (PGIC) score, PHQ-9 score, Patient-Reported Outcomes Measurement Information System–Ability to Participate in Social Roles and Activities (PROMIS-APS) score,[27] Short-Form 6-Dimension (SF-6D) score,[28] Short-Form 12-Dimension (SF-12) physical component summary score, SF-12 mental component summary score, and Somatic Symptom Disorder–B Criteria Scale (SSD-12) score[29] were reported by participants.

Outcomes

The primary outcome was self-reported PHQ-15 score 52 weeks after randomisation, analysed by intention to treat. The timing of this outcome permitted us to differentiate between immediate effects, which could have been non-specific results of extra attention, and sustained effects, which would have suggested ongoing benefit after cessation of the intervention, during analyses.

Secondary outcomes were quality of life (ie, via EQ-5D-5L, SF-6D, and ICECAP-A), depressive symptoms (ie, via PHQ-9), anxiety symptoms (ie, via General Anxiety Disorder-7 [GAD-7]),[30] ability to participate in social roles and activities (ie, via PROMIS-APS), PGIC, and health-care use. Health-care use was assessed via a self-report form to obtain data during the 52-week period in both primary and secondary care (appendix pp 137–140). We also included the HLS-EU6 measure of health literacy and the SSD-12 measure of the psychological burden of physical symptoms.

We considered potential harms in terms of adverse events (with a particular focus on delayed or emerging diagnoses of serious medical conditions), withdrawals in the intervention group indicating dissatisfaction, and missed opportunities for timely diagnosis through a reduction in diagnostic testing behaviour by GPs of participants in the intervention group. Adverse events were mainly identified through participant-reported outcome measures during scheduled collection of outcomes, but GPs with an extended role were required to report adverse events to CM, the trial manager, if they were detected during symptom-clinic consultations as well. Classification of serious versus non-serious adverse events is described in the protocol (appendix pp 84–86)

Statistical analysis

This trial was powered on a mean between-group difference of two points on the PHQ-15, with a pooled SD of five points.[31] This effect size was similar to effect sizes in two smaller European studies of extended consultations.[32,33] Calculations were made with 90% power and a 5% α level, and were inflated for 25% loss to follow-up and 6% potential treatment-centre imbalances or differences. The initial target sample size was 376 participants, but this target size was reduced to 350 participants during the final 6 months of the trial after discussion with the trial steering committee and the funder to maintain power at 90%, as the loss to follow-up at that time was less than anticipated during trial design. No interim analyses of trial data were done.

Outcomes were analysed with partly nested, heteroscedastic, mixed-effects regression models.[34,35] These models included fixed effects for baseline participant-reported outcome measure scores, gender, age, and allocation and random effects for the GP with an extended role in the intervention group only. Although stratification was by centre, we restricted clustering in the primary analysis to GP with an extended role because they treated participants from different centres. Analyses followed the intention-to-treat principle and no adjustments were made for multiplicity. Variation in the effectiveness of GPs with extended roles was quantified with the intraclass correlation coefficient from regression models.

Prespecified sensitivity analyses were conducted for the primary outcome. First, the main model was repeated excluding participants who had not completed their final questionnaire within 2 weeks before and 1 month after the 52-week timepoint. Second, the main model was conducted with a fixed study-centre effect. Third, to investigate potential COVID-19-pandemic effects, the allocation factor was entered as usual care, in-person intervention, or remote intervention. Fourth, missing data were imputed with multivariate imputation by chained equations (MICE). Finally, complier average causal effects (CACE) analysis was used to estimate the intervention effect among people who had the intervention allocated to them in either group via the per-protocol definition while preserving original allocation. CACE were modelled with two-stage, least-squares regression with compliance as an endogenous regressor instrumented by allocation. Baseline participant-reported outcome measure scores were used alongside gender and age as predictors in both MICE and CACE models. These models initially included GP with an extended role as a covariate. However, this inclusion resulted in statistical models that did not converge due to very small observed cluster effects. Numbers needed to treat to benefit were calculated post hoc via the reciprocal of the difference between the probabilities of a change of

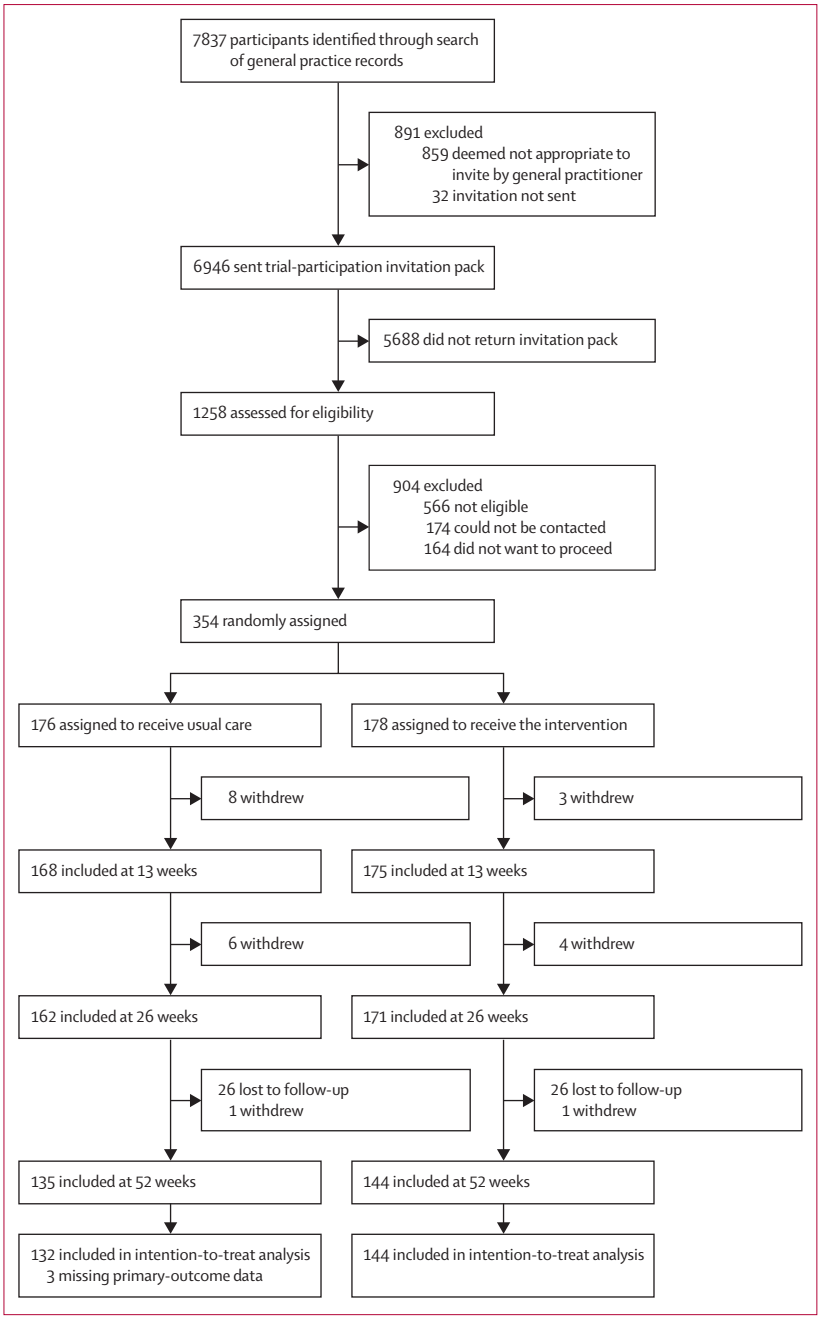

Figure 2: **Trial profile**

each magnitude; these effect sizes were multiples of minimal clinically important differences (MCIDs) for change from baseline in both groups from logistic-regression models adjusted for GP with an extended role in the intervention group only, baseline PHQ-15, age, gender, and effects of the intervention. We conducted an additional post-hoc analysis to map data from PHQ-15 and SSD-12 to estimate proportions of participants meeting criteria for somatic symptom disorder and bodily distress syndrome (appendix pp 5–6).

The trial was registered on the ISRCTN registry (ISRCTN57050216) on Oct 2, 2018, and the protocol has been published;[31] the statistical analysis plan was approved before data collection was completed and before the statistician was unmasked (appendix pp 24–60). All analyses were done in SAS version 9.4 and R version 4.2.1.

Role of the funding source

The funder of the trial reviewed the study protocol and could suggest changes, monitored trial progress through regular reports, and contributed to and approved the decision to reduce the target sample size from 376 to 350 because the rate of completion of the primary outcome was greater than originally predicted. The funder had no role in data collection, data analysis, data interpretation, or writing of the report.

Results

Between Dec 6, 2018, and Dec 21, 2021, 7837 participants were identified through a search of general practice electronic health records and 6946 were sent a trial-participation invitation pack via 108 GP practices (total estimated list size 913 331). After 1258 participants were assessed for eligibility, 904 were excluded, and 354 were enrolled and randomly assigned; 178 (50%) were assigned to receive the community-based symptoms clinic plus usual care and 176 (50%) were assigned to receive usual care only (figure 2). 276 (78%) of 354 participants were included in the primary-outcome analysis (132 [75%] of 176 receiving usual care only *vs* 144 [81%] of 178 receiving the intervention) as 52 (15%) were lost to follow-up, 23 (7%) withdrew, and three (1%) had missing data for the primary outcome.

Comparison of baseline data between participants included in the final analysis and those lost to follow-up did not contain any clinically important differences (appendix pp 16–18). To deliver the intervention, we initially recruited and trained seven GPs. However, two dropped out early on, one before seeing any participants, because of other commitments, so five GPs with an extended role provided the intervention.

Of the 354 participants, 291 (82%) identified as women and 63 (18%) identified as men. 25 (7%) identified as a minoritised ethnic group. 136 (38%) reported no academic qualifications after age 16, whereas 119 (34%) had at least one university degree (table 1; appendix pp 7–9). According to the HLS-EU6 survey, at baseline, only 122 (35%) of 354 participants met the cutoff for "sufficient" health literacy, 167 (47%) had scores in the range suggesting "problematic" health literacy, and 50 (14%) had "inadequate" health literacy (appendix pp 7–9). 240 participants (68%) were in paid full-time or part-time employment at baseline and 22 (6%) described themselves as able to work but not currently working. 268 (76%) of 354 participants were recruited from Yorkshire and the Humber, 46 (13%) were recruited from

Newcastle and Gateshead, 22 (6%) were recruited from northwest London, and 18 (5%) were recruited from Greater Manchester.

Symptoms reported as most troublesome via the free-text content of the PSYCHLOPS measure were recorded (appendix p 11),[22] as was the proportion of participants rating each symptom in the PHQ-15 as causing no bother, a little bother, or a lot of bother at baseline and their associations with SSD-12 (appendix pp 10, 12). 279 (79%) of 354 participants had moderate to high symptom-related distress on the SSD-12 measure (ie, total score ≥20).[29] We mapped symptoms to the four symptom groups from the bodily distress syndrome checklist (ie, autonomic, gastrointestinal, musculoskeletal, and general).[36] 274 participants (77%) had symptoms from all four groups, 78 (22%) had symptoms from three groups, and the remaining two participants (1%) had symptoms from two groups (appendix p 12).

Of the 178 participants who were randomly assigned to the intervention group, 165 (93%) attended the initial consultation, 156 (88%) attended the initial consultation plus at least one follow-up session, 143 (80%) attended the initial consultation plus at least two follow-up sessions, and 122 (69%) attended all sessions. None of the 13 participants who attended only one follow-up session completed the second PSYCHLOPS measure but 13 (62%) of the 21 participants who stopped after two follow-up sessions completed it. There were 24 unmasking incidents—15 (63%) were definite and nine (38%) were probable—across 816 participant outcome-data collections.

At baseline, PHQ-15 scores were 14·9 (SD 3·0) in the group receiving usual care and 15·0 (2·9) in the group receiving the intervention (figure 3). At the primary-outcome point of 52 weeks, PHQ-15 scores were 14·1 (3·7) in the group receiving usual care but had reduced to 12·2 (4·5) in the group receiving the intervention. The adjusted between-group difference of –1·82 (95% CI –2·67 to –0·97; table 2; appendix pp 19–21) was statistically significantly in favour of the intervention group (p<0·0001). PHQ-15 scores at 13 weeks were 14·1 (SD 4·0) in the group receiving usual care and 12·9 (4·1) in the group receiving the intervention. After adjustment for baseline participant-reported outcome measure scores, age, gender, and effects of the GP delivering the intervention, this between-group difference of –0·8 (95% CI –2·7 to 1·0) was not statistically significant. At 26 weeks, PHQ-15 scores in the group receiving usual care were 13·9 (SD 4·0) and in the group receiving the intervention were 12·7 (4·4), with an adjusted between-group difference of –1·0 (95% CI –2·3 to 0·2) that was not statistically significant. The primary effect estimate was consistent across all pre-agreed model specifications, including both the MICE and CACE models (appendix p 13). The intraclass correlation coefficient for GPs with an extended role showed little difference in effectiveness (<0·01).

	Usual care	Intervention	Total
Age, years			
Mean	46·1 (12·9)	45·2 (12·7)	45·6 (12·8)
Median	48·0 (36·8–56·0)	47·0 (35·2–56·0)	47·0 (36·0–56·0)
Range	20·0–69·0	18·0–70·0	18·0–70·0
Gender			
Man	32/176 (18%)	31/178 (17%)	63/354 (18%)
Woman	144/176 (82%)	147/178 (83%)	291/354 (82%)
Ethnicity			
Asian	7/176 (4%)	7/178 (4%)	14/354 (4%)
Mixed or other	6/176 (3%)	5/178 (3%)	11/354 (3%)
White	163/176 (93%)	166/178 (93%)	329/354 (93%)
First language			
English	163/176 (93%)	167/178 (94%)	330/354 (93%)
Other European language	3/176 (2%)	4/178 (2%)	7/354 (2%)
Asian language	5/176 (3%)	1/178 (1%)	6/354 (2%)
Other language	5/176 (3%)	6/178 (3%)	11/354 (3%)
Education			
Missing	3/176 (2%)	5/178 (3%)	8/354 (2%)
GSCE or equivalent	61/173 (35%)	59/173 (34%)	120/346 (35%)
A-level or equivalent	48/173 (28%)	43/173 (25%)	91/346 (26%)
Bachelor's degree	40/173 (23%)	49/173 (28%)	89/346 (26%)
Higher degree	15/173 (9%)	15/173 (9%)	30/346 (9%)
No formal qualifications	9/173 (5%)	7/173 (4%)	16/346 (5%)
PHQ-15 score			
Mean	14·9 (3·0)	15·0 (2·9)	14·9 (2·9)
Median	15·0 (12·9–17·0)	15·0 (12·2–17·0)	15·0 (12·9–17·0)
Range	10·0–21·9	10·0–23·1	10·0–23·1

Data are mean (SD), median (IQR), range, or n/N (%). PHQ-15=Patient Health Questionnaire-15.

Table 1: **Participant characteristics at baseline**

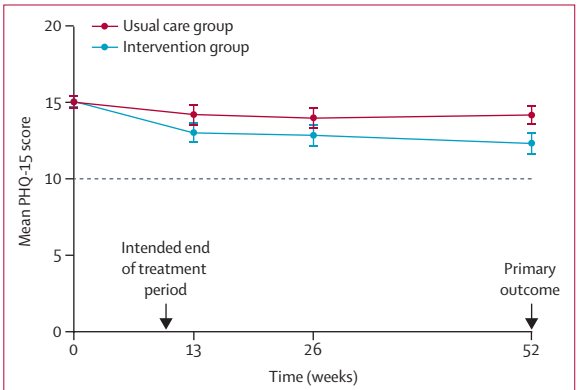

Figure 3: **PHQ-15 score at intervals from baseline by group**
Mean PHQ-15 score of 10 was the minimum threshold for inclusion in this study. PHQ-15=Patient Health Questionnaire-15.

Complete self-reported health-care use data were available for 166 (47%) of 354 participants, 77 in the intervention group and 89 in the group receiving usual care. There were only minor differences between groups in primary-care contacts, outpatient referrals, and diagnostic investigations (appendix pp 21–22).

	Usual care		Intervention		Adjusted* estimate of effect	Direction favours
	n	Mean	n	Mean		
PHQ-15	132	14·1 (3·7)	144	12·2 (4·5)	–1·82 (–2·67 to –0·97)	Intervention (p<0·0001)
EQ-5D-5L	128	0·498 (0·298)	130	0·572 (0·267)	0·072 (–0·001 to 0·145)	Intervention (p=0·054)
HLS-EU6	124	16·2 (4·0)	128	17·1 (3·9)	0·9 (0·1 to 1·8)	Intervention (p=0·032)
ICECAP-A	128	0·721 (0·210)	133	0·756 (0·195)	0·037 (–0·011 to 0·085)	Intervention (p=0·11)
PGIC	124	1·8 (1·2)	130	3·2 (1·9)	1·3 (0·7 to 1·9)†	Intervention (p=0·0019)
PHQ-9	133	10·9 (6·0)	142	8·8 (6·1)	–1·6 (–5·3 to 2·0)	Intervention (p=0·15)
PROMIS-APS	128	21·1 (8·6)	129	23·7 (9·0)	2·1 (0·0 to 4·2)	Intervention (p=0·053)
SF-6D	117	0·582 (0·109)	127	0·621 (0·139)	0·035 (–0·003 to 0·073)	Intervention (p=0·065)
SF-12 PCS score	111	36·1 (11·4)	124	37·1 (11·2)	0·9 (–1·9 to 3·7)	Intervention (p=0·49)
SF-12 MCS score	111	38·5 (11·7)	124	42·8 (11·5)	4·6 (1·7 to 7·4)	Intervention (p=0·0036)
SSD-12	128	22·9 (11)	133	21·6 (9·9)	–2·2 (–5·3 to 1·0)	Intervention (p=0·15)

Data are mean (SD), n, adjusted between-group difference, or p values. EQ-5D=5-Level EuroQol-5 Dimension. GP=general practitioner. HLS-EU6=European Health Literacy Survey-6. ICECAP-A=ICEpop Capability Measure for Adults. MCS=mental component summary. PCS=physical component summary. PGIC=Patient Global Impression of Change. PHQ-15=Patient Health Questionnaire-15. PROMIS-APS=Patient-Reported Outcomes Measurement Information System–Ability to Participate in Social Roles and Activities. SF-6D=Short-Form 6-Dimension. SF-12=12-Item Short Form Health Survey. SSD-12=Somatic Symptom Disorder–B Criteria Scale. *Adjusted for baseline participant-reported outcome measure scores, age, gender, allocation, and effects of the intervention GP with an extended role (in the intervention group only). †Unadjusted estimate presented as no baseline PGIC measure.

Table 2: **Estimates of intervention effect at 52 weeks via different participant-reported outcome measures**

There were 39 adverse events in the group receiving usual care and 36 adverse events in the group receiving the intervention (appendix p 22). There were no statistically significant between-group differences in the proportion of participants who had non-serious adverse events (–0·03, 95% CI –0·11 to 0·05) or serious adverse events (0·02, –0·02 to 0·07). The serious adverse events included one participant diagnosed with cancer, identified through routine screening; one participant with a history of irritable bowel syndrome who developed cholecystitis; and one participant who had a small pituitary adenoma found on imaging done due to dizziness. The most common adverse event was mental distress, occurring in 20 (11%) of 178 participants in the intervention group and 26 (15%) of 176 participants receiving usual care. No serious adverse event was deemed to be related to the trial intervention (eg, symptoms being mistakenly assumed not to represent the medical diagnosis, leading to the adverse event).

We calculated the probability distribution of change in PHQ-15 score from baseline to 52 weeks by allocation (appendix p 13). Post-hoc analysis indicated that the number needed to treat for a person to obtain a 2·3 point reduction (equivalent to the MCID) was 4·2 participants (95% CI 2·9–7·4), to obtain a 4·6 point reduction (equivalent to twice the MCID) was 4·9 participants (3·5–8·1), and to obtain a 6·9 point reduction (equivalent to three times the MCID) was 9·9 participants (7·0–17·0). Post-hoc subgroup analysis of the relationship between educational achievement and outcomes is shown in the appendix (p 15).

Discussion

Attendance at a community-based symptoms clinic with a GP with an extended role plus usual care led to a statistically significant improvement in the primary outcome of multiple and persistent physical symptoms compared with usual care only. The effect of this intervention, which lasted between 6 weeks and 12 weeks, was sustained 52 weeks after enrolment. There was no evidence of harms related to the intervention.

The trial recruited participants from four English regions, two of which are in areas of substantial socioeconomic disadvantage (ie, Yorkshire and the Humber and Newcastle and Gateshead). The choice of locations was made to produce findings that were broadly generalisable. The small number of people from minoritised ethnic groups was at least partly due to a need to restrict eligibility to people who were able to consult in English, without the need for translation. Although the number of participants included in this trial was small compared with the number of participants who were invited, we had anticipated this occurrence during trial design. Low enrolment rates are commonly observed when postal invitations are made from general practice electronic health-record searches, particularly when the search strategy favours sensitivity instead of specificity, as was the case in our trial. We chose this approach to reduce potential selection effects that might occur due to clinicians directly approaching participants.

The relatively large proportion of participants with no academic qualifications after age 16 (39%) was reflected in health literacy shown on the HLS-EU6 survey; at baseline, only 36% of participants met the cutoff for "sufficient" health literacy, 49% had scores in the range suggesting "problematic" health literacy, and 15% had "inadequate" health literacy.[25] Nevertheless, there were high rates of intervention completion and low rates of withdrawal from the study. None of the participants who

attended only one follow-up session completed the second PSYCHLOPS measure, but 62% of participants who stopped after two follow-up sessions completed it, suggesting that, for them, this was a planned conclusion. 78% of participants were included in the primary-outcome analysis, with only a small difference between groups (75% for usual care only *vs* 81% for the intervention group). Although responses to postal and telephone outcome-data collection could have differed, we regarded any possible bias introduced by this procedure as less important than the value of maximising retention in the study.

Delivery of the intervention was rigorously monitored and evaluated with good fidelity.[19] However, there were several limitations. For example, the symptom-clinic intervention was delivered by a small number of GPs with an extended role. However, none of these GPs had worked with the investigators before and they were selected by open competition. Although the GPs varied in experience and in confidence, at least initially, analysis showed no significant differences between them regarding outcomes. Supervision of the GPs with an extended role was by CB and VD, who developed the symptom-clinic intervention, and training of other GPs to conduct this intervention will need to be addressed in further work.

Although we excluded people with organic disease in the design of this trial, there is increasing evidence that persistent physical symptoms in the presence of long-term medical conditions share similar mechanisms to those addressed here,[13] so the model could be suitable in other situations. GPs with an extended role saw participants in research clinics that were not conducted in their own practices. They thus had no contact with study participants before or after the intervention. Unlike conventional time-restricted GP consultations in the UK, GPs with an extended role spent substantial time actively listening to the participant describe their illness and proposing and negotiating explanations and actions. Furthermore, the PHQ-15 has excellent internal reliability and good convergent validity with other measures of functioning, symptom severity, and disability days.[20]

Several limitations arose due to the pragmatic nature of the intervention and this trial. Participant inclusion criteria were deliberately broad in order to reflect the idea of persistent physical symptoms as an inclusive concept.[6] All participants had multiple and persistent physical symptoms across several different symptom groups. There was no attempt to conceal allocation from participants. However, all assessments were collected and processed with full concealment from trial team members. As this trial was pragmatic, there was no attention-control group, so we cannot exclude a nocebo effect of recruitment to the group receiving usual care only. However, the deliberate choice of primary outcome at 52 weeks, approximately 9–10 months after last contact with the symptom clinic, minimised the risk that self-reported outcomes reflected either the non-specific effects of attention or social desirability bias. The intervention was originally implemented as an in-person intervention without physical examination. However, during the COVID-19 pandemic, changing to remote online delivery of the intervention was necessary. Subgroup analysis showed that this change had little effect on outcomes. Although unplanned, this alteration suggests that the intervention can be delivered either in-person or remotely. Changes in access to both primary and secondary health care due to the pandemic are also likely to have led to reduced rates of consultation, referral, and testing for at least some symptoms. Although in typical circumstances up to half of referrals result in no diagnosis,[1] this proportion might have been less during our trial. We investigated wider societal costs as part of the sensitivity analyses conducted in the health economic evaluation (eg, including private health-care treatment costs and the costs of productivity losses due to illness) and these will be reported elsewhere. Furthermore, the original plan was to extract health-care use data from GP records, with participant self-report as a back-up. However, access to GP surgeries was severely limited by the COVID-19 pandemic, so self-report data were used in all analyses.

Two reviews of primary-care interventions for people with persistent physical symptoms published since 2020 found no evidence for effective primary care-based interventions.[37,38] The treatment model and findings of the process evaluation in this trial align with the desirable characteristics of a primary-care intervention outlined in the 2020 realist review.[38] A 2023 cluster-randomised trial of a shorter GP-delivered intervention focusing on return to work showed short-term benefits, but these benefits did not extend beyond 13 weeks of follow-up.[39] In MSS3, we observed changes in our primary outcome that were at least as large as those in a 2023 trial of transdiagnostic cognitive behavioural therapy in secondary care.[40] Although the mean treatment effect that we observed was less than the published clinically important difference of 2·3 points,[23] the 95% CIs around the point estimate included this value. Furthermore, the skewed distribution of treatment effect (appendix p 13) suggested that some participants benefitted particularly highly, such that the number needed to treat for twice the clinically important difference was five.

The findings of this trial lead to two questions about implementation. The first is whether the use of GPs with an extended role is appropriate when there is a shortage of GPs. The second is whether this type of intervention should be integrated into ordinary general practice. In relation to the first question, we see being a GP with an extended role as enhancing the core generalist skills of GPs, enabling them to focus on a particular group of people. Rather than diverting GPs from applying their core skills, the intervention provided GPs who expressed interest with an opportunity to develop expertise that they viewed as having a positive effect on both their

professional identity and everyday practice.[19] As initiatives to increase the number of GPs and supporting health professionals in primary care are enacted in the UK, a symptom-clinic service could be commissioned and delivered at local-area levels, such as through primary-care networks. In relation to the second question, we think that further work is necessary before integrating our intervention into ordinary general practice. Although the key elements of REAL appear to be relevant to short consultations, the knowledge and skills taught and learnt by GPs with an extended role in this trial took considerable time, practice, and supervision.[19] Further work is needed to find ways to condense elements of this intervention for use by GPs in ordinary clinical settings. This trial indicates that when appropriate criteria are used, the risk of previously unrecognised pathology in people with multiple and persistent physical symptoms is low enough that an approach of explaining and managing symptoms, while always considering other possible causes,[6] is appropriate.

A community-based symptoms clinic with a GP with an extended role that is focused on explanation and action led to a statistically significant reduction in multiple and persistent physical symptoms. This effect persisted at 9–10 months after the intervention. Future work should evaluate the cost-effectiveness of the intervention and find ways to identify people who are likely to benefit the most.

Contributors

CB and VD conceived the trial. CB, DW, JD, ARN, GR, ST, MH, MG, TS, RET, CC, and VD designed the trial and acquired funding. CM and JW collected data. KF conducted and analysed participant and general-practitioner interviews. CM, ET, and JW cleaned and validated the data. CB, GR, and WW were principal investigators for their local sites. CB and VD supervised the general practitioners with an extended role. LS, ARN, and JD accessed and verified the data and conducted the analysis. LS, ARN, JD, CB, and VD reviewed and interpreted the findings. EM provided a patient and public perspective throughout the trial design, delivery, and dissemination. CB drafted the original manuscript with support from LS, JD, DW, ARN, and CM. All authors reviewed and approved the final manuscript, had full access to all the data, and had final responsibility for the decision to submit for publication.

Declaration of interests

CB has received publisher royalties from Wiley. JD was a member of the Health and Social Care Delivery Research Board for the UK National Institute for Health and Care Research. All other authors declare no competing interests.

Data sharing

De-identified datasets and statistical code will be available with publication from the Clinical Trials Research Unit at the University of Sheffield upon request to ctruqa@sheffield.ac.uk. Requests should state the fields required and purpose of the request—ideally with a protocol but, at a minimum, with a research plan. The statistical analysis plan is available in the appendix (pp 24–60). The data dictionary can be made available on request. Requests will be considered on a case-by-case basis and requestors will be asked to complete a data-sharing agreement with the University of Sheffield before data transfer. Data will be retained for 6 years after June 30, 2023, before being destroyed.

Acknowledgments

This trial was funded by the UK National Institute for Health and Care Research (NIHR) Health Services and Delivery Research programme (project 15/136/07). The views expressed in this publication are those of the authors and not necessarily those of the National Health Service, the NIHR, or the UK Department of Health. We thank the members of the independent trial steering committee (Richard Morriss, Catherine Hewitt, Carolyn Chew-Graham, Laura Bojke, and Liza Barnett-Tweedie) and the data monitoring committee (David Kessler, Stefan Priebe, and Daisy Gaunt).

References

1 Nimnuan C, Hotopf M, Wessely S. Medically unexplained symptoms: an epidemiological study in seven specialities. *J Psychosom Res* 2001; **51:** 361–67.
2 Verhaak PF, Meijer SA, Visser AP, Wolters G. Persistent presentation of medically unexplained symptoms in general practice. *Fam Pract* 2006; **23:** 414–20.
3 Bermingham SL, Cohen A, Hague J, Parsonage M. The cost of somatisation among the working-age population in England for the year 2008–2009. *Ment Health Fam Med* 2010; **7:** 71–84.
4 Tomenson B, Essau C, Jacobi F, et al. Total somatic symptom score as a predictor of health outcome in somatic symptom disorders. *Br J Psychiatry* 2013; **203:** 373–80.
5 Burton C, Fink P, Henningsen P, Löwe B, Rief W. Functional somatic disorders: discussion paper for a new common classification for research and clinical use. *BMC Med* 2020; **18:** 34.
6 Rosendal M, Olde Hartman TC, Aamland A, et al. "Medically unexplained" symptoms and symptom disorders in primary care: prognosis-based recognition and classification. *BMC Fam Pract* 2017; **18:** 18.
7 Marks EM, Hunter MS. Medically unexplained symptoms: an acceptable term? *Br J Pain* 2015; **9:** 109–14.
8 Johansen M-L, Risor MB. What is the problem with medically unexplained symptoms for GPs? A meta-synthesis of qualitative studies. *Patient Educ Couns* 2017; **100:** 647–54.
9 Treufeldt H, Burton C. Stigmatisation in medical encounters for persistent physical symptoms/functional disorders: scoping review and thematic synthesis. *Patient Educ Couns* 2024; **123:** 108198.
10 Creed F, Guthrie E, Fink P, et al. Is there a better term than "medically unexplained symptoms"? *J Psychosom Res* 2010; **68:** 5–8.
11 Chen WG, Schloesser D, Arensdorf AM, et al. The emerging science of interoception: sensing, integrating, interpreting, and regulating signals within the self. *Trends Neurosci* 2021; **44:** 3–16.
12 Henningsen P, Gündel H, Kop WJ, et al. Persistent physical symptoms as perceptual dysregulation: a neuropsychobehavioral model and its clinical implications. *Psychosom Med* 2018; **80:** 422–31.
13 Löwe B, Andresen V, Van den Bergh O, et al. Persistent somatic symptoms across diseases—from risk factors to modification: scientific framework and overarching protocol of the interdisciplinary SOMACROSS research unit (RU 5211). *BMJ Open* 2022; **12:** e057596.
14 Burton C, Lucassen P, Aamland A, Olde Hartman T. Explaining symptoms after negative tests: towards a rational explanation. *J R Soc Med* 2015; **108:** 84–88.
15 Rolfe A, Burton C. Reassurance after diagnostic testing with a low pretest probability of serious disease: systematic review and meta-analysis. *JAMA Intern Med* 2013; **173:** 407–16.
16 Morton L, Elliott A, Thomas R, Cleland J, Deary V, Burton C. Developmental study of treatment fidelity, safety and acceptability of a symptoms clinic intervention delivered by general practitioners to patients with multiple medically unexplained symptoms. *J Psychosom Res* 2016; **84:** 37–43.
17 Royal College of General Practitioners. RCGP guide to GP clinical extended roles. 2021. https://www.rcgp.org.uk/getmedia/8725a327-5eaf-4a4d-b8f7-c8efa8668764/RCGP-Guide-to-GP-clinical-extended-roles-June-21-v21.pdf (accessed May 11, 2024).
18 Morton L, Elliott A, Cleland J, Deary V, Burton C. A taxonomy of explanations in a general practitioner clinic for patients with persistent "medically unexplained" physical symptoms. *Patient Educ Couns* 2017; **100:** 224–30.
19 Fryer K, Sanders T, Greco M, Mooney C, Deary V, Burton C. Recognition, explanation, action, learning: teaching and delivery of a consultation model for persistent physical symptoms. *Patient Educ Couns* 2023; **115:** 107870.
20 Kroenke K, Spitzer RL, Williams JB. The PHQ-15: validity of a new measure for evaluating the severity of somatic symptoms. *Psychosom Med* 2002; **64:** 258–66.

21 Kroenke K, Spitzer RL, Williams JB. The PHQ-9: validity of a brief depression severity measure. *J Gen Intern Med* 2001; **16:** 606–13.

22 Ashworth M, Robinson SI, Godfrey E, et al. Measuring mental health outcomes in primary care: the psychometric properties of a new patient-generated outcome measure, 'PSYCHLOPS' ('psychological outcome profiles'). *Primary Care Mental Health* 2005; **3:** 261–70.

23 Toussaint A, Kroenke K, Baye F, Lourens S. Comparing the Patient Health Questionnaire-15 and the Somatic Symptom Scale-8 as measures of somatic symptom burden. *J Psychosom Res* 2017; **101:** 44–50.

24 Herdman M, Gudex C, Lloyd A, et al. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; **20:** 1727–36.

25 Pelikan JM, Ganahl K. Measuring health literacy in general populations: primary findings from the HLS-EU Consortium's health literacy assessment effort. *Stud Health Technol Inform* 2017; **240:** 34–59.

26 Al-Janabi H, Flynn TN, Coast J. Development of a self-report measure of capability wellbeing for adults: the ICECAP-A. *Qual Life Res* 2012; **21:** 167–76.

27 Hahn EA, Devellis RF, Bode RK, et al. Measuring social health in the patient-reported outcomes measurement information system (PROMIS): item bank development and testing. *Qual Life Res* 2010; **19:** 1035–44.

28 Brazier J, Roberts J, Deverill M. The estimation of a preference-based measure of health from the SF-36. *J Health Econ* 2002; **21:** 271–92.

29 Toussaint A, Löwe B, Brähler E, Jordan P. The Somatic Symptom Disorder-B Criteria Scale (SSD-12): factorial structure, validity and population-based norms. *J Psychosom Res* 2017; **97:** 9–17.

30 Spitzer RL, Kroenke K, Williams JB, Löwe B. A brief measure for assessing generalized anxiety disorder: the GAD-7. *Arch Intern Med* 2006; **166:** 1092–97.

31 Mooney C, White DA, Dawson J, et al. Study protocol for the Multiple Symptoms Study 3: a pragmatic, randomised controlled trial of a clinic for patients with persistent (medically unexplained) physical symptoms. *BMJ Open* 2022; **12:** e066511.

32 Larisch A, Schweickhardt A, Wirsching M, Fritzsche K. Psychosocial interventions for somatizing patients by the general practitioner: a randomized controlled trial. *J Psychosom Res* 2004; **57:** 507–14.

33 Blankenstein AH, van der Horst HE, Schilte AF, et al. Development and feasibility of a modified reattribution model for somatising patients, applied by their own general practitioners. *Patient Educ Couns* 2002: **47:** 229–35.

34 Flight L, Allison A, Dimairo M, Lee E, Mandefield L, Walters SJ. Recommendations for the analysis of individually randomised controlled trials with clustering in one arm—a case of continuous outcomes. *BMC Med Res Methodol* 2016; **16:** 165.

35 Candlish J, Teare MD, Dimairo M, Flight L, Mandefield L, Walters SJ. Appropriate statistical methods for analysing partially nested randomised controlled trials with continuous outcomes: a simulation study. *BMC Med Res Methodol* 2018; **18:** 105.

36 Budtz-Lilly A, Fink P, Ørnbøl E, et al. A new questionnaire to identify bodily distress in primary care: the 'BDS checklist'. *J Psychosom Res* 2015; **78:** 536–45.

37 Byrne AK, Scantlebury A, Jones K, Doherty L, Torgerson DJ. Communication interventions for medically unexplained symptom conditions in general practice: a systematic review and meta-analysis of randomised controlled trials. *PLoS One* 2022; **17:** e0277538.

38 Leaviss J, Davis S, Ren S, et al. Behavioural modification interventions for medically unexplained symptoms in primary care: systematic reviews and economic evaluation. *Health Technol Assess* 2020; **24:** 1–490.

39 Abrahamsen C, Reme SE, Wangen KR, Lindbæk M, Werner EL. The effects of a structured communication tool in patients with medically unexplained physical symptoms: a cluster randomized trial. *EClinicalMedicine* 2023; **65:** 102262.

40 Chalder T, Patel M, Hotopf M, et al. Efficacy of therapist-delivered transdiagnostic CBT for patients with persistent physical symptoms in secondary care: a randomised controlled trial. *Psychol Med* 2023; **53:** 486–96.