The knowledge base for WISE,[1] which aims to encourage a structured and patient-centered approach to self-management support, includes:

- *Whole systems perspective* - with self-management support integrated into patient, practitioner and service organisation levels.

- *Team-based professional training* including GPs, nurses and administrative staff and designed to be relevant to routine clinical work, and flexible enough to respond to the needs of individual practices.[2,3]

- *Core principles of effective self-management support,* including priority and agenda setting, patient-centered information and shared decision-making

- *Facilitating access to community-based self-management resources*, such as Chronic Disease Self-Management Programs and patient support groups.

Outcome measures

The trial had three primary and nine secondary outcome measures, all pre-specified. Primary outcomes were collected at all three time-points (baseline, 6 months and 12 months); secondary outcomes were collected at 6 months and 12 months only, excepting for General Health. Secondary outcomes were not collected at baseline in order to reduce questionnaire length and maximize response. General health was, however, collected to provide a general measure of baseline health status. All outcome measures were previously validated instruments, with the exception of the self-care activities scale (see below). Outcome measures varied widely in scale range and direction, therefore to aid interpretability we rescaled all outcomes to a common 0-to-100 scale with a positive score indicating a better outcome; an exception was the EQ-5D which, as a standard economic measure, was kept on its original scale.

For analysis purposes, the 6-month measures of the three primary outcomes were treated as secondary outcomes.

Primary Outcomes

Shared decision making

The WISE approach is designed to assist patients and professionals to come to a shared understanding of the optimal ways of managing long-term conditions taking into account the patient's preferred management strategies and priorities regarding outcomes. Shared decision making was measured using the 6-item short-form Health Care Climate Questionnaire, which measures autonomy support provided by professionals (*e.g.*, provision of choice, acknowledging the individual's perspective) [4-7]

Self-efficacy

Self-efficacy refers to a self perception of confidence to complete activities. Self-efficacy was measured using a 5-item scale of confidence to undertake chronic disease management developed for this purpose [8] and used in previous trials of the Chronic Disease Self-management programme][9] and the Expert Patients Programme [10].

Health related quality of life

The EQ-5D is a generic measure of health related quality of life that allows the generation of QALYs. The 5-item scale covers mobility, self care, usual activities, pain, anxiety and depression, each with three levels of severity and provides a utility value based on a population tariff, where zero equates to dead and one equates to full health [8;11].

Secondary outcomes

The first four scales below were taken from the Medical Outcomes Survey [8]. They were selected for this study as each captures a distinct aspect of physical and psychological health and all performed well in the national evaluation of the Expert Patients Programme, with all bar General Health being significantly improved under that intervention.

General health

The General Health measure is a single item, rated on a five-point scale ranging from Excellent to Poor.

Social/Role Limitations

Score based on four items relating to the effect of health problems on daily, social and recreational activities, rated on a five-point scale ranging from Not At All to Almost totally.

Energy and Vitality

Score based on five items relating to levels of energy, tiredness, etc, rated on 6-point scales from None of the Time to All of the Time.

Psychological well-being

Five items about aspects of emotional life (happiness, nervousness, feeling 'down', etc), rated on 6-point scales from None of the Time to All of the Time.

Patient Enablement

The Patient enablement Instrument (PEI) is a validated 6-item scale of patient-reported enablement with respect to health problems, resulting from their last consultation. Each item is rated on a three-point scale (much better; better; same or less) with an additional 'not applicable' option.[12]

Health behaviour

A literature search failed to identify a validated generic instrument for assessing patient self-care activities. We therefore modified an existing – although non-validated - scale for arthritis patients consisting of 11 types of self-care (for example, exercise, diet, information seeking, avoiding stress) [6][13]. We re-expressed the scale items in a more generic fashion where necessary and added a twelfth item on taking medication as prescribed. We also revised the response set (Yes/No) to be more nuanced (Very true/Partly true/Not true).

Condition specific quality of life

We included separate disease-specific quality of life measures for each of the three exemplar conditions. These secondary outcomes, however, would only be analysed if a significant main effect or interaction with condition group was found on the generic quality of life measure (EQ-5D).

Disease specific quality of life in COPD was measured using the Short Breathing Problems Questionnaire, a shortened 10-item version of the Breathing Problems Questionnaire which measures the effects of COPD on mobility, self care, social activities energy and mental health [14;15].

Disease specific quality of life in diabetes was measured with the Diabetes Quality of Life questionnaire (DQOL) [16], a 15-item scale measuring issues such as satisfaction with treatment, impact on function and relationships, and self care.
Disease specific quality of life in IBS was measured using the IBS-QOL, a 34-item scale with eight subscale scores, including dysphoria, interference with activity, body image, health worry, food avoidance, social reaction, sex, and relationships [17]. We used a shortened version with 16 items.

Moderators

Illness perceptions (6 and 12 months)

Illness representations are implicit mental models of illness, which are related to coping strategies and outcomes[18]. One of the best known measures of illness representations is the Illness Perceptions Questionnaire. This study uses the brief 8-item IPQ, excluding the causal scale[19;20].

Health literacy

Health literacy is the 'degree to which individuals have the capacity to obtain, process and understand basic health information and services needed to make appropriate health decisions'[21] [22] and may place a limit on the ability of patients to engage with self-management. We will use questions on the patient's ability to read and understand health literature.

Social capital (6 and 12 months) Research indicates that social capital may have an influence on care pathways and management of illness[23]. All three domains of social capital (social networks, neighbourhood attachment & civic participation) are significantly associated with poor measures of health status[24]. Social capital will be measured with 14 items from the Health Survey for England (2003)[25].[26]

Measures used in the economic evaluation

For the economic analysis, our measure of effectiveness was the QALY, a composite measure of health generated from the EQ-5D instrument (see above). To assess service utilization (baseline, 6 and 12 months) we used a modification of the service utilisation questionnaire from the national evaluation of the Expert Patients Programme [10], which measures primary health care (GP visits, pharmacy), community health and social care, secondary health care services, out of pocket costs and costs of lost productivity.

Sample size and power

Sample size calculations were made on the basis of data collected from the national evaluation of the Expert Patients Programme [10]. Although all three patients groups were combined in the primary analysis (see below), we powered the trial to detect a fairly small effect of the intervention on diabetes, COPD, and IBS separately. The Expert Patients Programme evaluation had found effect sizes on most health status outcomes of around 0.2. We therefore powered this study to detect that level of effect.
Data on outcomes from Expert Patients Programme evaluation had a range of intraclass correlation coefficients (ICCs) from 0.01 to 0.07. For the power calculations, we assumed an ICC of 0.05. Baseline to follow up correlations were set at 0.6, towards the lower end of those found in the Expert Patients Programme. On these assumptions, each arm of the trial required 18 practices, and 36 patients per practice, to achieve 80% power to detect an effect size of 0.2. To allow for attrition of practices (estimated to be around 10%), we aimed to recruit 20 practices into each arm of the trial. Allowing for a 40% patient recruitment rate and 25% attrition of patients across the life of the study (based on our previous trials), an invitation to participate to 80 patients per practice with each exemplar condition was expected to provide an average of 48 patients per practice at baseline, reducing to 36 patients at 12 months.
Forty practices represented almost all of the 43 eligible practices in the PCT, and for most practices 80 patients included a high proportion, or all, of the patients on the practice disease

register for the condition, although a few practices had substantially higher numbers. On this basis, we decided to revise the sampling frame and to instead conduct a 'whole population' study by inviting all practices and all eligible patients into the study.

Analysis plan

An a-priori analysis plan, laying down full details of the primary and secondary outcomes, covariates, treatment of missing values, other methods and sensitivity analyses, was drawn up prior to analysis of the RCT. We applied intention-to-treat principles: practices were analysed in the trial arm they were randomised to, regardless of compliance to the training; and patients were included regardless of whether or not they were exposed to the intervention (ie by visiting their practice during the study period). The trial analyst was kept blind to practice allocation until all analyses had been completed.

Treatment of missing data

Missing data values for variables at baseline were substituted using a chained-equation multiple imputation (MI) procedure: five MI datasets were generated.

Missing values for outcome variables at follow-up was not imputed. Rather, the potential biasing effects of missing outcome data were addressed through covariate adjustment. Simulation studies show that adjusting for covariates related to missingness at follow-up is at least as efficient as multiple imputation at controlling for bias, and has been recommended [27]. Covariate adjustment has the additional advantage of increasing the precision of the treatment effect estimate. We did not apply inverse probability weighting (where weights are assigned to cases to adjust for missingness): this would in effect be duplicating the covariate approach above.

Distributional tests

We examined the distributional properties of each outcome variable. For variables where skewness or kurtosis was >1.0 we used the bootstrap method (sampling with replacement, with 100 samples) to derive the estimate of standard error for the intervention effect. We did not do tests for normality because the large sample size made these likely to be significant even for small deviations from normality.

Analysis methods

Each outcome was subjected to analysis of covariance within a multilevel regression framework. A two-level mixed model was used to account for the clustering of patients within practices. Practice intercept was treated as a random effect but all other variables were fixed effects. Each analysis controlled for baseline values of the outcome (where available), the design factors (practice list size, deprivation, and contractual type), plus additional covariates (see below). Maximum likelihood estimation was applied in view of the varying practice sample sizes and robust estimates of variance were used. Each of the five imputed datasets was analysed separately and the results combined using standard rules [28].

Although the study was powered to detect an effect size of 0.2 on each exemplar condition, to maximise overall power and minimise multiple testing we tested for a treatment effect across all three condition groups combined and for an interaction between trial arm and condition group (controlled for the main effects of condition group). If the interaction term was non-significant ($p>0.05$) we assumed that the treatment effect did not differ between condition groups and therefore did no condition-specific testing. If the interaction term was significant, we tested the treatment effect for each separate condition group against zero. An alpha-level of 0.05 was used throughout.

Covariates in the analysis

The three design factors (ie the minimisation variables of practice size, Index of Multiple Deprivation and type of contract) were automatically included in all analyses, as was the baseline value of the outcome variable when available. We also included a pre-specified set of covariates potentially prognostic of patient response to the intervention: gender; age; number of self-reported conditions; baseline general health; education; number of GP visits in 6 months prior to baseline; and residential IMD. Variables were entered in the forms given in Table 1 in the main

paper. Finally, we added baseline covariates found (from multivariate logistic regression) to predict a missing response at 12 months. Variables predictive of missing data were condition, age, general health, residential IMD and home ownership. All bar home ownership were already in the model as pre-specified covariates.

Sensitivity analyses

Two sensitivity analyses were undertaken to assess the stability of the results to the model specification. The first assessed the robustness of the main analysis results to the covariates in the model. This analysis removed the pre-specified covariates and predictors of missingness but still included the outcome at baseline and design factors.

The second sensitivity analysis assessed whether results had been influenced by patient non-exposure to either treatment or condition. This repeated the analysis after excluding patients who (i) did not visit their GP during the study period, or (ii) did not self-report the exemplar condition they were selected on. To avoid biasing the subsequent analysis, patients from both the intervention and control groups meeting these criteria were excluded.

Results of sensitivity analysis

Sensitivity analysis altered the conclusions of a significance test for one outcome only: shared decision making at 6-month follow-up. For this outcome, a significant intervention effect became non-significant under the analysis of sensitivity to exposure ($p=0.1$). Also, the significant result from the main analysis (and under sensitivity to covariate choice) favoured the control group, represents a very small effect, and in our opinion is probably spurious given the pattern of results for the rest of the outcomes.

Also for this outcome, a non-significant interaction between the intervention effect and disease condition group became significant under both sensitivity analyses ($p<0.05$). In further analyses of the individual condition groups, the intervention effect was not significantly different from zero for any group ($p>0.05$). The significance of the interaction term appeared to be due to slightly higher mean shared decision making score for intervention patients with diabetes relative to controls, compared to slightly lower mean scores for intervention patients with COPD or IBS relative to controls.

Reference List

(1) Kennedy A, Rogers A, Bower P. Support for self care for patients with chronic disease. *BMJ* 2007; 335(7627):968-970.

(2) Sikorski C, Luppa M, Konig HH, van den Bussche H, Riedel-Heller S. Does GP training in depression care affect patient outcome? - A systematic review and meta-analysis. *BMC Health Services Research* 2012; 12(1):10.

(3) Dawda P, Jenkins R, Varnam R. Quality improvement in general practice. An inquiry into the quality of general practice in England. 2010. London, King's Fund.

(4) Williams GC, Freedman ZR, Deci EL. Supporting autonomy to motivate patients with diabetes for glucose control. *Diabetes Care* 1998; 21(10):1644-1651.

(5) Williams GC, McGregor HA, Zeldman A, Freedman ZR, Deci EL. Testing a self-determination theory process model for promoting glycemic control through diabetes self-management. *Health Psychology* 2004; 23(1):58-66.

(6) Williams GC, McGregor HA, King D, Nelson CC, Glasgow RE. Variation in perceived competence, glycemic control, and patient satisfaction: relationship to autonomy support from physicians. *Patient Education and Counseling* 2005; 57(1):39-45.

(7) Ludman E, Simon G, Rutter C, Bauer M, Unutzer J. A measure for assessing patient perception of provider support for self -management of bipolar disorder. *Bipolar disorders* 2002; 4:349-353.

(8) Lorig K, Stewart AL, Ritter P, Gonzalez VM, Laurent D, Lynch J. Outcome measures for health education and other health care interventions. Thousand Oaks: Sage Publications; 1996.

(9) Lorig K, Sobel DS, Stewart AL, Brown BW, Bandura A, Ritter P et al. Evidence suggesting that a chronic disease self-management program can improve health status while reducing hospitalization a randomized trial. *Medical Care* 1999; 37(1):5-14.

(10) Kennedy A, Reeves D, Bower P, Lee V, Middleton E, Richardson G et al. The effectiveness and cost effectiveness of a national lay led self care support programme for patients with long-term conditions: a pragmatic randomised controlled trial. *J Epidemiol Community Health* 2007; 61:254-261.

(11) Kind P. The EuroQol instrument: an index of health-related quality of life. In: Spilker B, editor. Quality of life and pharmacoeconomics in clinical trials. 2nd ed. Philadelphia: Lippincott-Raven; 1996.

(12) Howie JGR, Heaney DJ, Maxwell M, Walker J. A comparison of a Patient Enablement Instrument (PEI) against two established satisfaction scales as an outcome measure of primary care consultations. *Fam Pract* 1998; 15(165):171.

(13) McDonald-Miszczak L, Wister AV. Predicting self-care behaviors among older adults coping with arthritis - A cross-sectional and 1-year longitudinal comparative analysis. *Journal of Aging and Health* 2005; 17(6):836-857.

(14) Hyland M, Bott J, Singh S, Kenyon C. Domains, constructs, and the development of the breathing problems questionnaire. *Quality of Life Research* 1994; 3:245-256.

(15) Hyland M, Singh S, Sodergren S, Morgan M. Development of a shortened version of the Breathing Problems Questionnaire suitable for use in a pulmonary rehabilitation clinic: a purpose-speciifc, disease-specific questionnaire. *Quality of Life Research* 1998; 7:227-233.

(16) Burroughs T, Desikan R, Waterman B, Gilin D, McGill J. Development and validation of the diabetes quality of life brief clinical inventory. *Diabetes Spectr* 2004; 17:41-49.

(17) Patrick DL, Drossman DA, Frederick IO, Dicesare J, Puder KL. Quality of life in Persons with Irritable Bowel Syndrome. *Digestive Diseases and Sciences* 1998; 43(2):400-411.

(18) Hagger MS, Orbell S. A meta-analytic review of the common-sense model of illness representations. *Psychology & Health* 2003; 18(2):141-184.

(19) Broadbent E, Petrie KJ, Main J, Weinman J. The Brief Illness Perception Questionnaire. *Journal of Psychosomatic Research* 2006; 60(6):631-637.

(20) Weinman J, Petrie KJ, MossMorris R, Horne R. The illness perception questionnaire: A new method for assessing the cognitive representation of illness. *Psychology & Health* 1996; 11(3):431-445.

(21) Department of Health and Human Services (USA). Healthy People 2010. 2004.

(22) Department of Health and Human Services (USA). Healthy People 2010. 2004.

(23) Pescosolido BA, Levy J. The role of social networks in health, illness, disease and healing: the accepting present, the forgotten past, and the dangerous potential for a complacent future. In: Pescosolido BA, Levy J, editors. *Social Networks and Health*. New York: JAI; 2002. 3-28.

(24) Petrou S, Kupek E. Social capital and its relationship with measures of health status: evidence from the health survey for England 2003. *Health Economics* 2008; 17(127):143.

(25) Bajekal M, Purdon S. *Social capital and social exclusion: development of a condensed module for the Health Survey for England* . 2011. London, *http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4058955.pdf* .

(26) Bajekal M, Purdon S. Social capital and social exclusion: development of a condensed module for the Health Survey for England (http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4058955.pdf - accessed 2nd November 2011). 2011. London, National Centre for Social Research.

(27) Groenwold R, Donders A, Roes K, Harrell Jr F, Moons K. Dealing with missing outcome data in randomized trials and observational studies. *American Journal of Epidemiology* 2012; 175:210-217.

(28) Rubin DB. Multiple Imputation for Nonresponse in Surveys. New York: John Wiley and Sons; 2004.