The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-003968
Article Type:	Research
Date Submitted by the Author:	06-Sep-2013
Complete List of Authors:	Peters, Michele; University of Oxford, Nuffield Department of Population Health Crocker, Helen; University of Oxford, Nuffield Department of Population Health Jenkinson, Crispin; University of Oxford, Nuffield Department of Population Health Doll, Helen; Oxford Outcomes, Fitzpatrick, Ray; University of Oxford, Nuffield Department of Population Health
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Health policy
Keywords:	Patient-reported outcome measures, Long-term conditions, PRIMARY CARE, Cohort study

The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey

Michele Peters [1], Senior Research Scientist
Helen Crocker [1], Researcher
Crispin Jenkinson [1], Professor of Health Services Research
Helen Doll [2], Senior Statistician
Ray Fitzpatrick [1], Professor of Public Health and Primary Care

[1] Nuffield Department of Population Health, University of Oxford, Old Road Campus, Oxford OX3 7LF, UK
[2] Oxford Outcomes, Seacourt Tower, West Way, Oxford OX2 0JJ, UK

Corresponding author

Michele Peters michele.peters@dph.ox.ac.uk

Telephone: 01865 289428

Nuffield Department of Population Health, Health Services Research Unit, University of Oxford, Old Road Campus, Oxford OX3 7LF

Keywords: Patient-reported outcome measures, long-term conditions, cohort study, primary care

Word count: 3327

ABSTRACT

Objectives

To evaluate the feasibility of using patient-reported outcome measures (PROMs) for long-term conditions (LTCs) in primary care

Design

A cohort postal survey conducted from September 2010 to April 2012

Setting

Primary care practices (n=33) in London and the North-West of England

Participants

4484 patients with a diagnosis of asthma, chronic obstructive pulmonary disease (COPD), diabetes, epilepsy, heart failure or stroke were sent a survey at baseline

Main outcome measures

Response rates for the baseline and follow-up surveys, and the EQ-5D, a generic patient-reported outcome measure.

Results

The baseline survey achieved a response rate of 38.4% (n=1721/4484) and at follow-up 71.4% (n=1136/1590). Response rates varied by LTC. Little change was found in health-related quality of life for the total sample (-0.001 for the EQ-5D score and 0.12 for the EQ-5D Visual Analogue Scale) between patients responding to both the baseline and follow-up surveys.

Conclusions

The response rate to the baseline survey was similar to that of other general practice surveys. Current UK policy aims to assess health service performance in LTCs by means of using PROMs. It thus would be desirable to improve response rates by making the invitation to self-reports of health-related quality of life more engaging for patients. Results of the study raise questions about optimal indicators for LTCs and appropriate timelines for assessment.

ARTICLE SUMMARY (up to 3 bullet points for each heading

Focus

- Long-term conditions (LTCs) reduce people's quality of life and increase the use of health care services
- The NHS Outcomes Framework aims to enhance quality of life in LTCs as assessed by patient-reported outcome measures (PROMs), in particular the EQ-5D

Key messages

- This survey's response rate of 38.4% at baseline was similar to response rates in other general practice surveys
- In those responding, little change was found in health-related quality of life over the course of one year
- In view of current policy in England, the response rates and limited change in health-related quality of life raises questions about the optimal indicators of care and appropriate timelines for assessment

Strengths and limitations

- The main strength of the study is that it is the first empirical evidence on the use of PROMs for LTCs in primary care focusing on patients with a confirmed diagnosis
- Another strength was that the study provides evidence that patients with LTCs who have completed a PROM once, are likely to complete a follow-up
- Limitations include the response rate, some logistic problems with identifying eligible patients from the GP databases and the number of patients excluded by the practices.

INTRODUCTION

The prevalence of long-term or chronic conditions, including multi-morbidity, is increasing world-wide [1-5]. Long-term conditions (LTCs) usually reduce people's quality of life, particularly in the case of multi-morbidity [2, 3, 6]. People with limiting LTCs are the most intensive users of the most expensive health care services [2] and the majority of health expenditure is for people with chronic conditions [5].

Chronic conditions are increasingly becoming the main concern of health care systems [5, 7] and a current challenge is to evaluate models of care and to develop new models that are more cost effective [4]. The Innovative Care for Chronic Conditions (ICCC) Framework aims to address different levels of health care including that at the micro level focused on patients and families.[8] In England, improving health care outcomes for all is the primary purpose of the National Health Service (NHS) in England. [9] The NHS Outcomes Framework's [10] second domain for improvement is the enhancement of quality of life for people with LTCs as assessed by patient-reported outcome measures (PROMs).

Patients' views are complementary to those of clinicians and provide unique information into the humanity and effectiveness of health care.[11] Routinely collecting data on patients' views through PROMS provides an opportunity to help drive change in the organisation and delivery of health care. Since 2009, PROMs have been used on a routine and mandated basis to assess outcomes in four elective surgical procedures in the NHS in England (http://www.ic.nhs.uk/proms, accessed 21.3.13). The role of PROMs in LTCs is more challenging than in elective surgery. In elective surgery, the PROMs are used to help assess the effectiveness of a single, discrete procedure. By contrast, LTCs are complex to manage involving diverse service providers and interventions over long timelines. The purpose of this study was to assess the feasibility of using PROMs to monitor the quality of life of individuals with a range of LTCs in primary care.

METHODS

A cohort survey was conducted in primary care from September 2010 to April 2012, with baseline and follow-up data being collected one year apart. The LTCs included were asthma, chronic obstructive pulmonary disease (COPD), diabetes, epilepsy, heart failure and stroke. Ethics approval had been obtained through the National Research Ethics (NRES) Committee of the Isle of Wight, Portsmouth & South East Hampshire (now the NRES South-Central Committee) in March 2010 and R&D approval from 20 participating Primary Care Trusts (PCTs). The study was registered on the National Institute for Health Research (NIHR) portfolio (UKCRN ID: 8462).

Design

A cohort survey was conducted by post. The baseline survey was provided pre-packed to participating practices and was mailed by practice staff, accompanied by a covering letter from the general practitioner. The baseline survey was returned to the Oxford research team and the follow-up survey was sent from Oxford. Consent was implied by returning the baseline survey. If participants were willing to complete a follow-up questionnaire, they were asked to give consent and to provide the Oxford team with their contact details. All surveys were numbered. This served a dual purpose, firstly to identify the respondents' practice and secondly to match baseline and follow-

up responses. A 'thank you/reminder' letter was sent by the practices two weeks after the mailing of the baseline survey to all patients invited into the study. The follow-up reminder, also sent after two weeks, was sent by the research team and targeted at non-responders only. A Microsoft Access database was set up to manage the mailing and receipt of the surveys.

Setting

The study was conducted in 33 primary care practices in London (n=18) and the North-West of England (NW) (n=15). Practices were recruited with the support of PCT staff, research nurses, local Primary Care Research Networks and the Diabetes Research Network in London. For less prevalent LTCs a larger number of practices participated. The majority of practices covered 3 LTCs and one practice covered 2 LTCs. Power calculations determined that a total of 1050 patients (or 175 per LTC) were needed to detect, at two-sided $p<0.05$ with 80% power, both a 15% difference in response rates between two independent groups and a moderate effect size of 0.3 in PROM score change over time. Ten practices provided patients for asthma (5 in London and 5 in NW), 16 for COPD (8 in London and 8 in NW), 10 for diabetes (5 in London and 5 in NW), 23 for epilepsy (13 in London and 10 in NW), 20 for heart failure (11 in London and 9 in NW) and 19 for stroke (12 in London and 7 in NW). Practices varied in size: 12 were small (<5800 patients), 13 medium (5800-10,500 patients) and 8 large (>10,500 patients). A slightly larger number of practices were recruited from more deprived areas (Table 1).

Table 1: Number of practices per social deprivation quintile

	Quintile	Range (IMD rank 2010*)	London	NW	Total
Most deprived	**Q1**	1 – 6496	4	4	8
	Q2	6497 - 12992	4	6	10
	Q3	12993 - 19488	5	1	6
	Q4	19489 - 25984	4	2	6
Least deprived	**Q5**	25985 - 32482	1	2	3

* Indices of Multiple Deprivation 2010

Participants

Eligible patients were identified through an automatic and remote search of practices' clinical systems by a subcontracted IT company prior to the baseline survey. The search is specific to every clinical system of which a range is available. The search was developed only for the 3 most widely used clinical systems. The aim was to identify approximately 50% of the patients for each LTC in every practice by a selection based on odd or even months of patients' birthdays. The search was based on Read codes in line with the criteria used in the Quality and Outcomes Framework (QOF) with two exceptions: patients with diabetes needed to be 18 years of age and patients with transient ischaemic attack(s) (TIAs) were excluded from the stroke group. The number of patients extracted by the search was compared to QOF estimates to check the expected numbers of patients were being selected. Patients were included in the survey for one LTC only; if they had multiple LTCs they were included for the rarest condition.

A member of staff (usually a GP or a nurse) reviewed the list of eligible patients identified from clinical systems to exclude any patients who were not considered suitable to be invited into the study (particularly if participation in the survey might cause serious distress).

Patient-reported outcome measures (PROMs)

A generic PROM and appropriate disease-specific PROMs were included in the surveys, as well as standard demographics questions and a question on comorbidities. The PROMs were selected on the basis of their psychometric properties evaluated by review work (http://phi.uhce.ox.ac.uk/newpubs.php , accessed 13.12.12) and licensing agreements. This article presents the findings from the EQ-5D [12], the generic measure used. The EQ-5D is a measure of health status primarily designed to provide a single-index value which represents the utility of specific health states, i.e. how given health states are valued by the general population.[12] It takes approximately five minutes to complete and comprises five items, one each on mobility, self-care, usual activities, pain/discomfort and anxiety/depression, that are all scored on a three-point scale. A single-index value is calculated from the five items typically with a score range from 1 (perfect health) to 0 (death) although a small number of scores below 0 can be obtained indicating states worse than death. The EQ-5D also includes a Visual Analogue Scale (VAS), ranging from 0 'worst imaginable health state' to 100 'best imaginable health state'.

Analysis

Response rates were calculated overall and for each LTC at baseline and follow-up. Cross-tabulations, with chi-square tests, were used to assess differences in response rates at baseline, consent to be followed up and response rate at follow-up between categorical variables (participant LTC, gender, age, location, ethnicity), and t-tests were used to compare mean scores for continuous data (time since diagnosis, number of comorbidities). Analysis of variance was used to identify factors (type of LTC, practice or patient related-factors or mean baseline EQ-5D scores) that were significantly related to follow-up response rates. The level of significance was set at two-sided $p<0.05$. All the variables were entered into a logistic regression (no consent to follow-up=0 vs. consent for follow-up=1, and follow-up not completed=0 vs. follow-up completed=1). Changes in the EQ-5D utility and VAS scores for respondents to both the baseline and follow-up surveys were assessed with paired t-tests. No data imputation was performed so the number of patients in the analysis of the EQ-5D may be lower than the number of respondents to the survey.

RESULTS

Recruitment of participants

The total number of patients extracted from each search was compared to 50% of the total QOF estimate (by LTC) based on the 2009/2010 QOF prevalence rates (http://www.qof.ic.nhs.uk/search/). The QOF estimates, number of patients extracted from the search, the overall difference and range of difference between the practices are presented in Table 2. A positive score on the difference means that a higher number of patients than expected were extracted from the search, whereas a negative number means that a lower than expected number was extracted. Although these results give an indication of how well the searches worked, they need to be interpreted with caution. The number of patients extracted in small practices was low. The

extracted numbers for diabetes and stroke were expected to be lower than the number of patients in QOF. Diabetes patients aged 17 and upwards are included in QOF, whereas this study only included diabetes patients aged 18 or over. This difference is expected to be minimal. Stroke presented more of a challenge as QOF includes TIAs which were not included in this study. Therefore, it was unlikely that the numbers extracted from the search would closely match the QOF estimates.

Table 2: Numbers of eligible patients extracted in PROMs pilot search compared to QOF estimates

LTC	QOF estimate (n)	PROMs pilot search (n)	Difference		Range of difference for practices (%)	
			n	%		
Asthma	7722	7615	-107	-1.4	-48.1	+62.6
COPD	2036	1590	-446	-21.9	-71.2	+42.1
Diabetes	5878	5154	-824	-13.8	-28.6	+19.1
Epilepsy	822	1088	+266	-32.4	-80.0	+348.3
Heart failure	948	937	-11	-1.2	-29.4	+17.7
Stroke	2351	1071	-1280	-54.4	-71.2	-30.0

Exclusions by practices

The rates of patients excluded from being sent a questionnaire varied between LTCs and practices (Table 3). The rate of exclusions tended to be higher in the NW than in London, although this difference was small for asthma, COPD and diabetes. However, the difference was substantial for epilepsy (56.6% in NW vs. 10.0% in London), heart failure (28.5% NW vs. 12.9% London) and stroke (24.2% NW vs. 16.8% London). The reasons for excluding patients were death, blindness, co-morbidities (such as dementia or cerebral palsy) or learning difficulties, patients having moved or being on an extended holiday, not having a clear diagnosis, recently having a traumatic life event or being in a care home. Comorbidities and learning difficulties were particularly common reasons for excluding epilepsy patients.

Table 3: Exclusions of patients by practices (cohort baseline data)

LTC (n practices)	N patients extracted		N patients excluded		% included	
	Total	Practice range	Total	Practice range	Total	Practice range
Asthma (10)	1628	64-684	294	0-197	81.9	63.5-100.0
COPD (16)	602	8-88	35	6-80	94.2	81.5-100.0
Diabetes (10)	1169	63-185	48	0-13	95.9	90.2-100.0
Epilepsy (23)	985	4-260	460	4-78	53.3	19.2-100.0
Heart failure (20)	687	5-143	167	4-81	75.7	56.6-100.0
Stroke (19)	525	4-69	107	0-19	79.6	52.4-100.0

Response rates

Cohort baseline

A total of 4486 questionnaires were sent and 1721 were returned achieving an overall response rate of 38.4%. The response rate varied between LTCs, with the response rate being the highest in heart

failure (50.4%, n=262) and the lowest in asthma (30.0%, n=400) (Table 4). There was significant variation in response rates between LTCs (p<0.001), by practice (p=0.018) and by region (p=0.002). Additionally, across all LTCs there were significant non-linear relationships between response rates and the practices' deprivation score (p=0.024, Table 5), and the practices' EQ-5D mean score adjusted for patient age, gender, time since diagnosis and number of comorbidities (p=0.004). Practices with a QOF score of 100 (maximum score) had significantly (p=0.013) higher response rates (mean response rate=42.9%, SD 11.9) than those who did not (mean=35.5%, SD 14.3).

Table 4: Cohort baseline survey numbers sent and returned, and response rates (%) for each LTC (overall, by region and practice range).

LTC (n practices)	N		Overall RR (%)	Regional RR (%)		RR by practice (range) (%)
	Sent	Returned		London	NW	
Asthma (10)	1334	395	30.0	22.7	33.0	14.3-50.0
COPD (16)	567	279	49.2	43.0	54.3	32.1-66.7
Diabetes (10)	1121	448	40.0	30.5	50.1	28.1-61.3
Epilepsy (23)	525	180	34.0	35.5	33.6	0-53.9
Heart failure (20)	520	262	50.0	48.8	51.1	30.6-71.4
Stroke (19)	418	152	36.4	30.0	44.0	7.7-63.2

Table 5: Mean response rate by deprivation score (N refers to the number of practices for each LTC covered i.e. 97 cases)

Deprivation quintile	N	Mean	SD
Most deprived	23	34.2	13.5
2	29	43.6	13.5
3	18	39.7	10.2
4	18	43.2	9.8
Least deprived	9	48.0	14.1

Cohort follow-up

At baseline, 93.1% of responding patients (1603 of 1721) had agreed to be sent a follow-up questionnaire. The lowest rate of consent was achieved for heart failure (90.5%) and the highest for COPD (95.3%). Thirteen patients were excluded from follow-up as they indicated that they had not been diagnosed with the LTC referred to; the denominator for the follow-up response rate was thus 1590. For the total sample, there was no significant difference in whether consent to follow-up was given by LTC, age, region, ethnicity, number of comorbidities or time since diagnosis. There was, however, a difference between practices (p=0.008) with the proportion of baseline respondents giving consent ranging from 78.6% to 100%. There were some disease-specific differences in consent, including gender (96.1% of men consented vs. 90.1% of women, p=0.029) and practice (consent in asthma ranged from 79.1%-100% in 10 practices, p=0.010) in asthma, and number of comorbidities for epilepsy (mean 1.17, SD 1.30 for those who consented vs. mean 2.14, SD 2.00 for those who did not consent, p=0.017) and heart failure (mean 1.92, SD 1.62 for those who consented vs. mean 2.60, SD 1.71 for those who did not consent, p=0.049) respectively.

The overall response rate of baseline respondents who consented to the follow-up was 71.4% (n=1136/1590). Numbers of questionnaires sent and response rates by LTC are given in Table 6.

Differences in response rates to the follow-up survey were examined by means of baseline characteristics of respondents who had consented to the follow-up. For the overall sample, there were significant differences in response rate at follow-up by LTC (p=0.015), age (p<0.001), ethnicity (p=0.008) and region (p=0.007), with epilepsy and heart failure patients being less likely to respond than patients with one of the other LTCs, as were younger patients, those based in London and those from ethnic minority backgrounds (Table 7). The baseline mean EQ-5D score was significantly lower (p<0.001) in non-responders to follow-up (mean 0.66, SD 0.33) than in follow-up responders (mean 0.73, SD 0.29). There were no significant differences in the response rate at follow-up for gender, time since diagnosis, number of comorbidities or practice.

Table 6: Cohort follow-up survey numbers sent and returned, and response rates (%) for each LTC (overall, by region and practice range).

LTC (n practices)	N		Overall RR (%)	Regional RR (%)		RR by practice (range) (%)
	Sent	Returned		London	NW	
Asthma (10)	366	267	73.0	65.4	75.1	53.9-82.4
COPD (16)	262	187	71.4	68.9	73.1	44.4-82.9
Diabetes (10)	424	321	75.7	67.9	80.7	60.4-87.8
Epilepsy (23)	166	104	62.7	69.8	58.3	25.0-100.0
Heart failure (20)	234	155	66.2	60.6	68.7	25.0-100.0
Stroke (19)	137	102	74.5	70.5	77.6	0-100.0

Table 7: Factors significantly related to the number of questionnaires returned at cohort follow-up

		% responders
LTC (p=0.015)	Asthma	72.9
	COPD	71.4
	Diabetes	75.7
	Epilepsy	62.7
	Heart failure	66.2
	Stroke	74.5
Age (years) (p<0.001)	18-24	37.5
	25-34	48.4
	35-44	63.4
	45-54	70.1
	55-64	78.8
	75-84	71.7
	85+	65.2
Region (p=0.007)	London	67.2
	North-West	73.7
Ethnicity (p=0.008)	White	72.9
	Other	63.0

EQ-5D scores

EQ-5D scores could be calculated for 93.1% (n=1058) of the respondents to both the baseline and follow-up surveys. The EQ-5D scores and VAS were found to be significantly different between LTCs both at baseline and follow-up (all p<0.001). However, no significant differences were found for the mean EQ-5D scores for the total sample between baseline and follow-up. For the VAS, a significant

difference was found for heart failure only, with scores deteriorating at follow-up (mean difference -3.53, 95% CI -6.67 to -0.38). The data are presented in Table 8.

Table 8: EQ-5D and VAS mean, and change, scores at baseline and follow-up

		Asthma	COPD	Diabetes	Epilepsy	Heart failure	Stroke	Total
EQ-5D York tariff								
n		255	177	301	95	137	93	1058
Baseline	**Mean**	0.83	0.67	0.73	0.76	0.64	0.67	0.73
	95% CI	0.80-0.86	0.63-0.71	0.69-0.76	0.71-0.82	0.59-0.69	0.61-0.74	0.71-0.75
Follow-up	**Mean**	0.84	0.67	0.72	0.76	0.64	0.67	0.73
	95% CI	0.81-0.87	0.63-0.71	0.69-0.76	0.71-0.81	0.59-0.69	0.60-0.73	0.71-0.75
Difference	**Mean**	0.01	-0.002	-0.003	-0.001	-0.005	-0.008	-0.001
	95% CI	-0.004-0.015	-0.019-0.014	-0.013-0.008	-0.021-0.018	-0.022-0.013	-0.31-0.13	-0.007-0.045
p		NS	NS	NS	NS	NS	NS	NS
	EQ-5D VAS							
n		248	173	296	91	145	82	1035
Baseline	**Mean**	73.77	62.29	68.16	71.40	62.20	73.84	68.54
	95% CI	71.31-76.23	59.31-65.27	65.84-70.48	67.11-75.68	58.93-65.47	67.18-76.75	67.27-69.81
Follow-up	**Mean**	74.33	62.14	69.76	73.59	58.67	71.96	68.42
	95% CI	71.94-76.72	59.13-65.16	67.53-71.99	69.68-77.50	55.10-62.23	67.18-76.75	67.16-69.68
Difference	**Mean**	0.56	-0.15	1.60	2.20	-3.53	-1.88	0.12
	95% CI	-1.09-2.21	-2.94-2.66	-0.19-3.39	-0.96-5.36	-6.67--0.38	-5.12-1.37	-0.87-1.10
p		NS	NS	NS	NS	0.029	NS	NS

DISCUSSION

This study provides the first empirical evidence on the use of PROMs for LTCs in primary care focusing on patients with a confirmed diagnosis. The evidence on participation of individuals with LTCs is essential for making decisions about the feasibility of expanding the current PROMs programme to LTCs. The aim of the study was to evaluate the feasibility of collecting PROMs data in LTCs through primary care focusing on response rates to a baseline and follow-up survey conducted one year apart. Overall, a 38.4% response rate was achieved at baseline and 71.5% for the cohort (those responding at baseline and agreeing to be sent a follow-up questionnaire) at follow-up. Response rates varied between LTCs at both assessments. Other factors significantly related to response rates were practice factors (including deprivation score, QOF score and mean EQ-5D score) at baseline, and respondent factors (including age, ethnicity, region and mean baseline EQ-5D score) at follow-up. Other methods to assess feasibility included testing the remote and automatic search of GP databases for eligible patients and to monitor patient exclusions. Problems were encountered with the search. It was possible to overcome these to some extent by the IT company amending the searches. However, discrepancies remained in the number of patients extracted from the search when compared with QOF estimates. The PROMs programme aims to be as inclusive as possible and therefore the number of patients excluded by practices at baseline was monitored. The exclusion rates varied by LTC and by region. The variation by LTC may be related to the nature of the specific

LTC or the comorbidities associated with a particular LTC. The highest exclusion rate was for epilepsy, which practices reported to be expected as many people with epilepsy have learning difficulties or other comorbidities which would make it inappropriate for them to be included in the study. There were no significant changes in EQ-5D scores between baseline and follow-up. The VAS was significantly different between baseline and follow-up for heart failure only.

The main strength of this study is the potential for greater reliability of the diagnosis obtained from GP clinical systems over the self-reported diagnoses in the GP Patient Survey [13, 14]. While the study was able to detect patients with LTC using remote and automatic search of GP databases, weaknesses of the methods tested in this study were the high exclusion rates by practices in some LTCs and the logistics of remotely identifying eligible patients from GP databases. Commonly reported reasons for exclusion were comorbidities such as learning difficulties or mental health problems. This suggests that it may be not be feasible to collect PROMs data from all patients with some diagnoses; other methods, such as collecting data by proxy, may have to be considered. Furthermore, the problems found in identifying eligible patients from the practice databases means that further development work and testing would be necessary to ensure that such a system was feasible and able to be used across all clinical systems.

Response rates were the main variable of interest of this study, and it would be desirable for response rates, particularly those at baseline, to be higher. However, other general practice-based surveys such as the GP Patient Survey have achieved a similar response rate [13, 14], suggesting that such a low response rate would be typical. The advantage of this study is that the potential for greater reliability of the diagnosis obtained from GP clinical systems over the self-reported diagnoses in the GP Patient Survey. The response rates for the data collection of the four elective PROMs were higher (60.3%) than the baseline response rates reported here, as were the response rates post-surgery (87.5%) [15]. Overall response rates to health surveys are decreasing [16], but the magnitude of response rates are not always thought to be a problem. Roland and colleagues found little evidence that low response rates and non-response bias in health surveys led to unfairness in payments for performance in general practices [13]. It is not known whether there is any non-response bias in this study, and efforts to enhance response rates should not be dismissed. Traditional methods to increase response rates include a good cover letter and sending a reminder. More modern methods, however, such as using electronic surveys may help increase response rates, especially in younger age groups. In this respect, it is worth noting that the response rate at follow-up of 71.5% suggests that once patients are 'on-board' the study they are likely to continue to respond, even one year later.

The other variables of interest in this study were the EQ-5D mean and VAS scores reported by the survey respondents, with very little change occurring in health-related quality of life. In view of the goal of the current policy in England to assess performance of the health service by means of the EQ-5D, these results raise broader questions about optimal indicators of care for long-term conditions and appropriate timelines for assessment. Currently there are a number of major uncertainties regarding trends and timelines for the progression of LTCs. Moreover, given evidence from this study that better baseline EQ-5D scores resulted in higher response rates at follow-up, it is unclear how helpful self-reported health-related quality of life is in monitoring such trends.

To improve response rates, the invitation to contribute self-reports of health needs to be more engaging in the sense of patients perceiving that their reports will actually serve a purpose. This greater sense of point or purpose could be achieved in three distinct, but not mutually exclusive ways. Firstly, patients could find the information valuable and informative, for example by providing feedback of their progress over time or in comparison with other patients. Secondly, patients could find the information in PROMs helpful in preparing for consultations with healthcare providers or as a part of regular review or assessment. Thirdly, it is conceivable that patients would value providing information if it were truly the case that PROMs data were used to provide evidence of quality or performance of services. The NHS and indeed all other healthcare systems are a long way from being able to support any of these possible uses of PROMs for long-term conditions that might enhance patient engagement. Further research needs to address whether PROMs can better inform patients about their progress, support communication of need or facilitate contributions to quality assessment.

Acknowledgments

We would like to thank the Primary Care Research Networks in London and the North-West; the Diabetes Research Network in London; all the practice staff who helped with the recruitment of practices and patients, and all the patients who participated in the surveys. Furthermore, we would like to thank Mrs Sarah Dummett for her support in collecting the data.

Competing interest declaration

All authors have completed the Unified Competing Interest form at www.icmje.org./coi_disclosure.pdf (available on request from the corresponding author) and declare that (1) All authors have support from the Department of Health (England) for the submitted work; (2) none of the authors have a relationships with commercial companies that might have an interest in the submitted work in the previous 3 years; (3) their spouses, partners, or children have no relationships that may be relevant to the work; and (4) none of the authors have non-financial interests that may be relevant to the submitted work.

Details of contributions

RF, CJ, and HD conceived and designed the study. MP and HC collected data. All authors contributed to the data analysis. MP wrote the article, with contributions from HC, HD, CJ and RF. All authors had access to the data in the study, and can take responsibility for the integrity of the data and the accuracy of the data analysis. All authors edited and approved the final version of the manuscript.

Funding and sponsor

This research was initially funded as an individual project under the Department of Health's Policy Research Programme. From 2011, it was incorporated within the research programme of the Department of Health funded Policy Research Unit on Quality and Outcomes of Person Centred Care (QORU), a collaboration between the London School of Economics and Political Science (LSE) and the Universities of Kent and Oxford.

The University of Oxford Clinical Trials and Research Governance (CTRG) acted as sponsor.

Disclaimer

This is an independent report commissioned and funded by the Policy Research Programme in the Department of Health. The views expressed are not necessarily those of the Department.

Ethics approval

Ethics approval was given by the National Research Ethics (NRES) Committee of the Isle of Wight, Portsmouth & South East Hampshire (now the NRES South-Central Committee) in March 2010 (reference number 10/H0501/10). A change of the diabetes PROM was approved in September 2010.

For the baseline, consent was implied by completion and return of the survey. Consent to the follow-up was given by the majority of baseline participants by providing the research team with contact details as part of completing the baseline questionnaire.

Registration

This study was registered on the National Institute for Health Research (NIHR) portfolio in June 2010 (UKCRN ID: 8462).

REFERENCES

1. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012, **380**:37-43
2. Department of Health. Long Term Conditions Compendium of Information. Third ed. London: Department of Health, 2012.
3. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007;**22 Suppl 3**:391-5
4. Pearson WS, Bhat-Schelbert K, Probst JC. Multiple chronic conditions and the aging of America: A challenge for primary care physicians. Journal of Primary Care and Community Health 2012;**3**(1):51-56
5. Anderson G, Horvath J. The growing burden of chronic disease in America. Public Health Rep 2004;**119**(3):263-70
6. Heyworth IT, Hazell ML, Linehan MF, Frank TL. How do common chronic conditions affect health-related quality of life? Br J Gen Pract 2009;**59**(568):e353-8
7. Epping-Jordan JE, Pruitt SD, Bengoa R, Wagner EH. Improving the quality of health care for chronic conditions. Qual Saf Health Care 2004;**13**(4):299-305
8. World Health Organization. Innovative care for chronic conditions: building blocks for action. Geneva, 2002.
9. Department of Health. Equity and Excellence: Liberating the NHS. London, 2010.
10. Department of Health. The NHS Outcomes Framework 2012/2013. London, 2011.
11. Black N, Jenkinson C. How can patients' views of their care enhance quality improvement? BMJ 2009;**399**:202-05
12. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001;**33**(5):337-43
13. Roland M, Elliott M, Lyratzopoulos G, et al. Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. BMJ 2009;**339**:b3851
14. Campbell J, Smith P, Nissen S, Bower P, Elliott M, Roland M. The GP Patient Survey for use in primary care in the National Health Service in the UK--development and psychometric characteristics. BMC Fam Pract 2009;**10**:57
15. Browne J, Jamieson L, Lewsey J, et al. Patient reported outcome measures (PROMs) in elective surgery. Report to the Department of Health, 2007.
16. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007;**17**(9):643-53

STROBE Statement—checklist of items that should be included in reports of observational studies

	Item No	Recommendation	
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	✓
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	✓
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	✓
Objectives	3	State specific objectives, including any prespecified hypotheses	✓
Methods			
Study design	4	Present key elements of study design early in the paper	✓
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	✓
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	✓
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	✓
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	✓
Bias	9	Describe any efforts to address potential sources of bias	N/A
Study size	10	Explain how the study size was arrived at	✓
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	✓
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	✓
		(*b*) Describe any methods used to examine subgroups and interactions	✓
		(*c*) Explain how missing data were addressed	✓
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	✓
		(*e*) Describe any sensitivity analyses	

Continued on next page

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	✓
		(b) Give reasons for non-participation at each stage	✓
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	✓
		(b) Indicate number of participants with missing data for each variable of interest	✓
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	✓
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	✓
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	✓
Discussion			
Key results	18	Summarise key results with reference to study objectives	✓
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	✓
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	✓
Generalisability	21	Discuss the generalisability (external validity) of the study results	N/A
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	✓

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.

The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2013-003968.R1
Article Type:	Research
Date Submitted by the Author:	15-Jan-2014
Complete List of Authors:	Peters, Michele; University of Oxford, Nuffield Department of Population Health Crocker, Helen; University of Oxford, Nuffield Department of Population Health Jenkinson, Crispin; University of Oxford, Nuffield Department of Population Health Doll, Helen; Oxford Outcomes, Fitzpatrick, Ray; University of Oxford, Nuffield Department of Population Health
Primary Subject Heading:	Health services research
Secondary Subject Heading:	Health policy
Keywords:	Patient-reported outcome measures, Long-term conditions, PRIMARY CARE, Cohort study

The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey

Michele Peters [1], Senior Research Scientist
Helen Crocker [1], Researcher
Crispin Jenkinson [1], Professor of Health Services Research
Helen Doll [2], Senior Statistician
Ray Fitzpatrick [1], Professor of Public Health and Primary Care

[1] Nuffield Department of Population Health, University of Oxford, Old Road Campus, Oxford OX3 7LF, UK
[2] Oxford Outcomes, Seacourt Tower, West Way, Oxford OX2 0JJ, UK

Corresponding author

Michele Peters michele.peters@dph.ox.ac.uk

Telephone: 01865 289428

Nuffield Department of Population Health, Health Services Research Unit, University of Oxford, Old Road Campus, Oxford OX3 7LF

Keywords: Patient-reported outcome measures, long-term conditions, cohort study, primary care

Word count: 3327

ABSTRACT

Objectives

To evaluate the feasibility of using patient-reported outcome measures (PROMs) for long-term conditions (LTCs) in primary care

Design

A cohort postal survey conducted from September 2010 to April 2012

Setting

Primary care practices (n=33) in London and the North-West of England

Participants

4484 patients with a diagnosis of asthma, chronic obstructive pulmonary disease (COPD), diabetes, epilepsy, heart failure or stroke were sent a survey at baseline

Main outcome

The main outcome was to evaluate the feasibility of and the recruitment strategies for collecting PROMs data in LTCs by assessing the response rates for the baseline and follow-up surveys.

Secondary outcomes were the evaluation of change scores of the EQ-5D index and visual analogue scale (VAS) between baseline and follow-up surveys

Results

The baseline survey achieved a response rate of 38.4% (n=1721/4485) and at follow-up 71.5% (n=1136/1589). Response rates varied by LTC. Little change was found in health-related quality of life for the total sample (-0.001 for the EQ-5D index score and 0.12 for the EQ-5D VAS) between patients responding to both the baseline and follow-up surveys.

Conclusions

The response rate to the baseline survey was similar to that of other general practice surveys. Current UK policy aims to assess health service performance in LTCs by means of using PROMs. It thus would be desirable to improve response rates by making the invitation to self-reports of health-related quality of life more engaging for patients. Results on the EQ-5D score raise questions about optimal indicators for LTCs and appropriate timelines for assessment.

ARTICLE SUMMARY

Focus

- Long-term conditions (LTCs) reduce people's quality of life and increase the use of health care services
- The NHS Outcomes Framework aims to enhance quality of life in LTCs as assessed by patient-reported outcome measures (PROMs), in particular the EQ-5D

Key messages

- This survey's response rate of 38.4% at baseline was similar to response rates in other general practice surveys
- In those responding, little change was found in health-related quality of life over the course of one year
- In view of current policy in England, the response rates and limited change in health-related quality of life raises questions about the optimal indicators of care and appropriate timelines for assessment

Strengths and limitations

- The main strength of the study is that it is the first empirical evidence on the use of PROMs for LTCs in primary care focusing on patients with a confirmed diagnosis
- Another strength is that the study provides evidence that patients with LTCs who have completed a PROM once, are likely to complete a follow-up
- Limitations include the response rate, some logistic problems with identifying eligible patients from the GP databases and the number of patients excluded by the practices.

INTRODUCTION

The prevalence of long-term or chronic conditions, including multi-morbidity, is increasing world-wide [1-5]. Long-term conditions (LTCs) usually reduce people's quality of life, particularly in the case of multi-morbidity [2, 3, 6]. People with limiting LTCs are the most intensive users of the most expensive health care services [2] and the majority of health expenditure is for people with chronic conditions [5].

Chronic conditions are increasingly becoming the main concern of health care systems [5, 7] and a current challenge is to evaluate models of care and to develop new models that are more cost effective [4]. The Innovative Care for Chronic Conditions (ICCC) Framework aims to address different levels of health care including that at the micro level focused on patients and families.[8] In England, improving health care outcomes for all is the primary purpose of the National Health Service (NHS) in England [9] and the NHS Outcomes Framework [10] sets out the indicators for measuring health outcomes in NHS services. The Outcomes Framework's second domain for improvement is the enhancement of quality of life for people with LTCs as assessed by patient-reported outcome measures (PROMs).

Patients' views are complementary to those of clinicians and provide unique information into the humanity and effectiveness of health care.[11] Routinely collecting data on patients' views through PROMS provides an opportunity to help drive change in the organisation and delivery of health care. Since 2009, PROMs have been used on a routine and mandated basis to assess outcomes in four elective surgical procedures in the NHS in England (http://www.ic.nhs.uk/proms, accessed 21.3.13). The role of PROMs in LTCs is more challenging than in elective surgery. In elective surgery, the PROMs are used to help assess the effectiveness of a single, discrete procedure. By contrast, LTCs are complex to manage involving diverse service providers and interventions over long timelines. The purpose of this study was to assess the feasibility of using PROMs to monitor aggregated quality of life data for a range of LTCs in primary care.

METHODS

A cohort survey was conducted in primary care from September 2010 to April 2012, with baseline and follow-up data being collected one year apart. The primary aim of the study was to evaluate the recruitment strategies for and the feasibility of collecting PROMs data in LTCs by assessing response rates at two points of data collection. Secondary outcomes were the change scores of the EQ-5D index and visual analogue scale (VAS) between baseline and follow-up surveys. Two considerations informed the secondary outcomes of the study: on the one hand the NHS Outcomes Framework's [12] second domain refers to the possibility of changes over time in the quality of life in LTCs and on the other hand, the trajectory of many LTCs is deterioration in health status. The LTCs included were asthma, chronic obstructive pulmonary disease (COPD), diabetes, epilepsy, heart failure and stroke. Ethics approval had been obtained through the National Research Ethics (NRES) Committee of the Isle of Wight, Portsmouth & South East Hampshire (now the NRES South-Central Committee) in March 2010 and R&D approval from 20 participating Primary Care Trusts (PCTs). The study was registered on the National Institute for Health Research (NIHR) portfolio (UKCRN ID: 8462).

Design

A cohort survey was conducted by post. The baseline survey was provided pre-packed to participating practices and was mailed by practice staff, accompanied by a covering letter from the general practitioner. The baseline survey was returned to the Oxford research team and the follow-up survey was sent from Oxford. Consent was implied by returning the baseline survey. If participants were willing to complete a follow-up questionnaire, they were asked to give consent and to provide the Oxford team with their contact details. All surveys were numbered. This served a dual purpose, firstly to identify the respondents' practice and secondly to match baseline and follow-up responses. A 'thank you/reminder' letter was sent by the practices two weeks after the mailing of the baseline survey to all patients invited into the study. The follow-up reminder, also sent after two weeks, was sent by the research team and targeted at non-responders only. A Microsoft Access database was set up to manage the mailing and receipt of the surveys.

Setting

The study was conducted in 33 primary care practices in London (n=18) and the North-West of England (NW) (n=15). Practices were recruited with the support of PCT staff, research nurses, local Primary Care Research Networks and the Diabetes Research Network in London. For less prevalent LTCs a larger number of practices participated. The majority of practices covered 3 LTCs and one practice covered 2 LTCs. Power calculations determined that a total of 1050 patients (or 175 per LTC) were needed to detect, at two-sided $p<0.05$ with 80% power, both a 15% difference in response rates between two independent groups and a moderate effect size of 0.3 in PROM score change over time. Ten practices provided patients for asthma (5 in London and 5 in NW), 16 for COPD (8 in London and 8 in NW), 10 for diabetes (5 in London and 5 in NW), 23 for epilepsy (13 in London and 10 in NW), 20 for heart failure (11 in London and 9 in NW) and 19 for stroke (12 in London and 7 in NW). The number of practices per LTC varied due to the prevalence of the LTC and practice size. Practices varied in size: 12 were small (<5800 patients), 13 medium (5800-10,500 patients) and 8 large (>10,500 patients). A slightly larger number of practices were recruited from more deprived areas (Table 1).

Table 1: Number of practices per social deprivation quintile

	Quintile	Range (IMD rank 2010*)	London	NW	Total
Most deprived	**Q1**	1 – 6496	4	4	8
	Q2	6497 - 12992	4	6	10
	Q3	12993 - 19488	5	1	6
	Q4	19489 - 25984	4	2	6
Least deprived	**Q5**	25985 - 32482	1	2	3

* Indices of Multiple Deprivation 2010

Participants

Eligible patients were identified through an automatic and remote search of practices' clinical systems by a subcontracted IT company prior to the baseline survey. The search is specific to every clinical system of which a range is available. The search was developed only for the 3 most widely used clinical systems. The aim was to identify approximately 50% of the patients for each LTC in every practice by a selection based on odd or even months of patients' birthdays. The search was based on Read codes in line with the criteria used in the Quality and Outcomes Framework (QOF)

with two exceptions: patients with diabetes needed to be 18 years of age and patients with transient ischaemic attack(s) (TIAs) were excluded from the stroke group. The number of patients extracted by the search was compared to QOF estimates to check the expected numbers of patients were being selected. Patients were included in the survey for one LTC only; if they had multiple LTCs they were included for the rarest condition.

A member of staff (usually a GP or a nurse) reviewed the list of eligible patients identified from clinical systems to exclude any patients who were not considered suitable to be invited into the study. The instruction to practices was to exclude patients if participation in the survey might cause serious distress. The practice staff made the judgement of when this criterion applied.

Patient-reported outcome measures (PROMs)

A generic PROM and appropriate disease-specific PROMs were included in the surveys, as well as standard demographics questions and a question on comorbidities. The PROMs were selected on the basis of their psychometric properties evaluated by review work (http://phi.uhce.ox.ac.uk/newpubs.php , accessed 13.12.12) and licensing agreements. This article presents the findings from the EQ-5D [13], the generic measure used. The EQ-5D is a measure of health status primarily designed to provide a single-index value which represents the utility of specific health states, i.e. how given health states are valued by the general population.[13] It takes approximately five minutes to complete and comprises five items, one each on mobility, self-care, usual activities, pain/discomfort and anxiety/depression, that are all scored on a three-point scale. A single-index value is calculated from the five items typically with a score range from 1 (perfect health) to 0 (death) although a small number of scores below 0 can be obtained indicating states worse than death. The EQ-5D also includes a Visual Analogue Scale (VAS), ranging from 0 'worst imaginable health state' to 100 'best imaginable health state'.

Analysis

Response rates were calculated overall and for each LTC at baseline and follow-up. Cross-tabulations, with chi-square tests, were used to assess differences in response rates at baseline, consent to be followed up and response rate at follow-up between categorical variables (participant LTC, gender, age, location, ethnicity), and t-tests were used to compare mean scores for continuous data (time since diagnosis, number of comorbidities). Analysis of variance was used to identify factors (type of LTC, practice or patient related-factors or mean baseline EQ-5D scores) that were significantly related to follow-up response rates. All the variables relevant to cohort follow-up response rates were entered into a logistic regression (no consent to follow-up=0 vs. consent for follow-up=1, and follow-up not completed=0 vs. follow-up completed=1). The level of significance was set at two-sided $p<0.05$.

Changes in the EQ-5D utility and VAS scores for respondents to both the baseline and follow-up surveys were assessed with paired t-tests. No data imputation was performed so the number of patients in the analysis of the EQ-5D may be lower than the number of respondents to the survey. The level of significance was set at $p<0.05$.

RESULTS

Recruitment of participants

The total number of patients extracted from each search was compared to 50% of the total QOF estimate (by LTC) based on the 2009/2010 QOF prevalence rates (http://www.qof.ic.nhs.uk/search/). The number of patients extracted by the searches is higher (total n=17,455), than those invited into the study (n=5596), as the search data was available for all LTCs for every practice but surveys were sent only to up to three LTCs per practice. The flow of the 5596 patients invited into the study is illustrated in Figure 1. The QOF estimates, number of patients extracted from the search the overall difference and range of difference between the practices are presented in Table 2. A positive score on the difference means that a higher number of patients than expected were extracted from the search, whereas a negative number means that a lower than expected number was extracted. Although these results give an indication of how well the searches worked, they need to be interpreted with caution. The number of patients extracted in small practices was low. The extracted numbers for diabetes and stroke were expected to be lower than the number of patients in QOF. Diabetes patients aged 17 and upwards are included in QOF, whereas this study only included diabetes patients aged 18 or over. This difference is expected to be minimal. Stroke presented more of a challenge as QOF includes TIAs which were not included in this study. Therefore, it was unlikely that the numbers extracted from the search would closely match the QOF estimates.

Table 2: Numbers of eligible patients extracted in PROMs pilot search compared to QOF estimates

LTC	QOF estimate (n)	PROMs pilot search (n)	Difference		Range of difference for practices (%)	
			n	%		
Asthma	7722	7615	-107	-1.4	-48.1	+62.6
COPD	2036	1590	-446	-21.9	-71.2	+42.1
Diabetes	5878	5154	-824	-13.8	-28.6	+19.1
Epilepsy	822	1088	+266	+32.4	-80.0	+348.3
Heart failure	948	937	-11	-1.2	-29.4	+17.7
Stroke	2351	1071	-1280	-54.4	-71.2	-30.0

Exclusions by practices

The rates of patients excluded from being sent a questionnaire varied between LTCs and practices (Table 3). The rate of exclusions tended to be higher in the NW than in London, although this difference was small for asthma, COPD and diabetes. However, the difference was substantial for epilepsy (56.6% in NW vs. 10.0% in London), heart failure (28.5% NW vs. 12.9% London) and stroke (24.2% NW vs. 16.8% London). The reasons for excluding patients were death, blindness, co-morbidities (such as dementia or cerebral palsy) or learning difficulties, patients having moved or being on an extended holiday, not having a clear diagnosis, recently having a traumatic life event or being in a care home. Comorbidities and learning difficulties were particularly common reasons for excluding epilepsy patients.

Table 3: Exclusions of patients by practices (cohort baseline data)

LTC (n practices)	N patients extracted		N patients excluded		% included	
	Total	Practice range	Total	Practice range	Total	Practice range
Asthma (10)	1628	64-684	294	0-197	81.9	63.5-100.0

COPD (16)	602	8-88	35	6-80	94.2	81.5-100.0
Diabetes (10)	1169	63-185	48	0-13	95.9	90.2-100.0
Epilepsy (23)	985	4-260	460	4-78	53.3	19.2-100.0
Heart failure (20)	687	5-143	167	4-81	75.7	56.6-100.0
Stroke (19)	525	4-69	107	0-19	79.6	52.4-100.0
Total (33)	5596	38-880	1111	0-261	80.1	35.5-100.0

Response rates

Cohort baseline

A total of 4485 questionnaires were sent and 1721 were returned achieving an overall response rate of 38.4%. Thirteen patients were excluded from analysis and follow-up as they reported not to have been diagnosed with the specified LTC. The response rate varied between LTCs, with the response rate being the highest in heart failure (50.4%, n=262) and the lowest in asthma (30.0%, n=400) (Table 4). There was significant variation in response rates between LTCs (p<0.001), by practice (p=0.018) and by region (p=0.002). Additionally, across all LTCs there were significant non-linear relationships between response rates and the practices' deprivation score (p=0.024, Table 5), and the practices' EQ-5D mean score adjusted for patient age, gender, time since diagnosis and number of comorbidities (p=0.004). Practices with a QOF score of 100 (maximum score) had significantly (p=0.013) higher response rates (mean response rate=42.9%, SD 11.9) than those who did not (mean=35.5%, SD 14.3).

Table 4: Cohort baseline survey numbers sent and returned, and response rates (%) for each LTC (overall, by region and practice range).

LTC (n practices)	N		Overall RR (%)	Regional RR (%)		RR by practice (range) (%)
	Sent	Returned		London	NW	
Asthma (10)	1334	400	30.0	22.7	33.0	14.3-50.0
COPD (16)	567	279	49.2	43.0	54.3	32.1-66.7
Diabetes (10)	1121	448	40.0	30.5	50.1	28.1-61.3
Epilepsy (23)	525	180	34.0	35.5	33.6	0-53.9
Heart failure (20)	520	262	50.0	48.8	51.1	30.6-71.4
Stroke (19)	418	152	36.4	30.0	44.0	7.7-63.2
TOTAL (33)	4485	1721	38.4	32.7	42.2	14.2-58.9

Table 5: Mean response rate by deprivation score (N refers to the number of practices for each LTC covered i.e. 98 cases)

Deprivation quintile	N	Mean	SD
Most deprived	23	34.2	13.5
2	30	43.6	13.5
3	18	39.7	10.2
4	18	43.2	9.8
Least deprived	9	48.0	14.1

Cohort follow-up

At baseline, 92.3% of responding patients (1589 of 1721) had agreed and were eligible to be sent a follow-up questionnaire. The lowest rate of consent was achieved for heart failure (90.5%) and the highest for COPD (95.3%). For the total sample, there was no significant difference in whether consent to follow-up was given by LTC, age, region, ethnicity, number of comorbidities or time since diagnosis. There was, however, a difference between practices (p=0.008) with the proportion of baseline respondents giving consent ranging from 78.6% to 100%. There were some differences in consent for individual LTCs, including gender (96.1% of men consented vs. 90.1% of women, p=0.029) and practice (consent in asthma ranged from 79.1%-100% in 10 practices, p=0.010) in asthma, and number of comorbidities for epilepsy (mean 1.17, SD 1.30 for those who consented vs. mean 2.14, SD 2.00 for those who did not consent, p=0.017) and heart failure (mean 1.92, SD 1.62 for those who consented vs. mean 2.60, SD 1.71 for those who did not consent, p=0.049) respectively.

The overall response rate of baseline respondents who consented to the follow-up was 71.5% (n=1136/1589). Numbers of questionnaires sent and response rates by LTC are given in Table 6. Differences in response rates to the follow-up survey were examined by means of baseline characteristics of respondents who had consented to the follow-up. For the overall sample, there were significant differences in response rate at follow-up by LTC (p=0.015), age (p<0.001), ethnicity (p=0.008) and region (p=0.007), with epilepsy and heart failure patients being less likely to respond than patients with one of the other LTCs, as were younger patients, those based in London and those from ethnic minority backgrounds (Table 7). The baseline mean EQ-5D score was significantly lower (p<0.001) in non-responders to follow-up (mean 0.66, SD 0.33) than in follow-up responders (mean 0.73, SD 0.29). There were no significant differences in the response rate at follow-up for gender, time since diagnosis, number of comorbidities or practice.

Table 6: Cohort follow-up survey numbers sent and returned, and response rates (%) for each LTC (overall, by region and practice range).

LTC (n practices)	N		Overall RR (%)	Regional RR (%)		RR by practice (range) (%)
	Sent	Returned		London	NW	
Asthma (10)	366	267	73.0	65.4	75.1	53.9-82.4
COPD (16)	262	187	71.4	68.9	73.1	44.4-82.9
Diabetes (10)	424	321	75.7	67.9	80.7	60.4-87.8
Epilepsy (23)	166	104	62.7	69.8	58.3	25.0-100.0
Heart failure (20)	234	155	66.2	60.6	68.7	25.0-100.0
Stroke (19)	137	102	74.5	70.5	77.6	0-100.0
Total (33)	1589	1136	71.4	67.3	73.7	53.3-66.2

Table 7: Factors significantly related to the number of questionnaires returned at cohort follow-up

		% responders
LTC (p=0.015)	Asthma	73.0
	COPD	71.4
	Diabetes	75.7
	Epilepsy	62.7
	Heart failure	66.2
	Stroke	74.5
Age (years) (p<0.001)	18-24	37.5
	25-34	48.4

	35-44	63.4
	45-54	70.1
	55-64	78.8
	75-84	71.7
	85+	65.2
Region (p=0.007)	London	67.3
	North-West	73.7
Ethnicity (p=0.008)	White	72.9
	Other	63.0

When analyses were performed of prediction of return of the follow-up questionnaire for each of the six LTCs, significant differences were observed for some factors, including age (p<0.001) and ethnicity (p=0.009) in asthma; age (p=0.012) in COPD, gender (p=0.032) and region (p=0.003) in diabetes; age (p=0.011) in epilepsy; and ethnicity (p=0.003) in stroke. The direction of these differences was the same as for the overall sample (i.e. lower completion rate in younger patients and ethnic minorities), with the additional finding that in diabetes, women were less likely to complete the follow-up than men. When entered into a logistic regression (follow-up not completed =0 vs. follow-up completed =1), only some age groups and EQ5D remained significant (**Table 8**). Asthma and age 18-24 served as reference categories.

Table 8: Factors related to completion of the follow-up questionnaire

		p	Odds ratio
LTC	**Asthma**	NS	--
	COPD	NS	1.07
	Diabetes	NS	0.80
	Epilepsy	NS	1.06
	Heart Failure	NS	0.75
	Stroke	NS	0.66
Age (years)	**18-24**	<0.001	--
	25-34	0.001	0.24
	35-44	0.005	0.37
	45-54	NS	0.71
	55-64	NS	1.04
	65-74	NS	1.68
	75-84	0.035	1.70
	85+	NS	1.23
	Region	NS	1.20
	EQ5D	<0.001	2.11
	Ethnicity	NS	0.77
	Gender	NS	0.88
	Constant	NS	1.32

EQ-5D scores

EQ-5D scores could be calculated for 93.1% (n=1058) of the respondents to both the baseline and follow-up surveys. The EQ-5D scores and VAS were found to be significantly different between LTCs both at baseline and follow-up (all p<0.001). However, no significant differences were found for the

mean EQ-5D scores for the total sample between baseline and follow-up. For the VAS, a significant difference was found for heart failure only, with scores deteriorating at follow-up (mean difference -3.53, 95% CI -6.67 to -0.38). The data are presented in Table 9.

Table 9: EQ-5D and VAS mean, and change, scores at baseline and follow-up

		Asthma	COPD	Diabetes	Epilepsy	Heart failure	Stroke	Total
EQ-5D York tariff								
n		255	177	301	95	137	93	1058
Baseline	**Mean**	0.83	0.67	0.73	0.76	0.64	0.67	0.73
	95% CI	0.80-0.86	0.63-0.71	0.69-0.76	0.71-0.82	0.59-0.69	0.61-0.74	0.71-0.75
Follow-up	**Mean**	0.84	0.67	0.72	0.76	0.64	0.67	0.73
	95% CI	0.81-0.87	0.63-0.71	0.69-0.76	0.71-0.81	0.59-0.69	0.60-0.73	0.71-0.75
Difference	**Mean**	0.01	-0.002	-0.003	-0.001	-0.005	-0.008	-0.001
	95% CI	-0.004-0.015	-0.019-0.014	-0.013-0.008	-0.021-0.018	-0.022-0.013	-0.31-0.13	-0.007-0.045
p		NS	NS	NS	NS	NS	NS	NS
		EQ-5D VAS						
n		248	173	296	91	145	82	1035
Baseline	**Mean**	73.77	62.29	68.16	71.40	62.20	73.84	68.54
	95% CI	71.31-76.23	59.31-65.27	65.84-70.48	67.11-75.68	58.93-65.47	67.18-76.75	67.27-69.81
Follow-up	**Mean**	74.33	62.14	69.76	73.59	58.67	71.96	68.42
	95% CI	71.94-76.72	59.13-65.16	67.53-71.99	69.68-77.50	55.10-62.23	67.18-76.75	67.16-69.68
Difference	**Mean**	0.56	-0.15	1.60	2.20	-3.53	-1.88	0.12
	95% CI	-1.09-2.21	-2.94-2.66	-0.19-3.39	-0.96-5.36	-6.67--0.38	-5.12-1.37	-0.87-1.10
p		NS	NS	NS	NS	0.029	NS	NS

DISCUSSION

This study provides the first empirical evidence on the use of PROMs for LTCs in primary care focusing on patients with a confirmed diagnosis. The evidence on participation of individuals with LTCs is essential for making decisions about the feasibility of expanding the current PROMs programme to LTCs. The aim of the study was to evaluate the feasibility of collecting PROMs data in LTCs through primary care focusing on response rates to a baseline and follow-up survey conducted one year apart. Overall, a 38.4% response rate was achieved at baseline and 71.5% for the cohort (those responding at baseline and agreeing to be sent a follow-up questionnaire) at follow-up. Response rates varied between LTCs at both assessments. Other factors significantly related to response rates were practice factors (including deprivation score, QOF score and mean EQ-5D score) at baseline, and respondent factors (including age, ethnicity, region and mean baseline EQ-5D score) at follow-up. Other methods to assess feasibility included testing the remote and automatic search of GP databases for eligible patients and to monitor patient exclusions. Problems were encountered with the search. It was possible to overcome these to some extent by the IT company amending the searches. However, discrepancies remained in the number of patients extracted from the search when compared with QOF estimates. The PROMs programme aims to be as inclusive as possible and therefore the number of patients excluded by practices at baseline was monitored. The exclusion

rates varied by LTC and by region. The variation by LTC may be related to the nature of the specific LTC or the comorbidities associated with a particular LTC. The highest exclusion rate was for epilepsy, which practices reported to be expected as many people with epilepsy have learning difficulties or other comorbidities which would make it inappropriate for them to be included in the study. There were no significant changes in EQ-5D scores between baseline and follow-up. The VAS was significantly different between baseline and follow-up for heart failure only.

The main strength of this study is the potential for greater reliability of the diagnosis obtained from GP clinical systems over the self-reported diagnoses in the GP Patient Survey [14, 15]. While the study was able to detect patients with LTC using remote and automatic search of GP databases, weaknesses of the methods tested in this study were the lack of information about non-responders of the baseline survey, high exclusion rates by practices in some LTCs and the logistics of remotely identifying eligible patients from GP databases. No information about non-responders had been collected as there were concerns that collecting these data was too burdensome for participating practices. However, such data would have been valuable to assess non-response bias beyond the non-response bias from the follow-up. Exclusions rates varied between practices. This was partly due to problems with identifying eligible patients in the practices where the search by the IT company was first conducted. For example, in the case of epilepsy the initial search did not include the code in relation to medication (as by QOF) and some practices chose for a member of staff to exclude ineligible patients rather than the search being re-run. Some differences in exclusions occurred due to practices' different interpretations of the instruction to exclude patients in whom invitation into the study would cause serious distress, although this is believed to only have had a minor impact. Reasons for exclusions were not recorded systematically (again to minimize the burden on practices). Practices informally reported that the most common reasons for exclusion were comorbidities such as learning difficulties or mental health problems. This suggests that it may be not be feasible to collect PROMs data from all patients with some diagnoses; other methods, such as collecting data by proxy, may have to be considered. Remotely identifying eligible patients was challenging due to mistakes made in the initial searches (search criteria for asthma and epilepsy did not meet QOF criteria) and the search working less well in some of the clinical systems used by the participating practices. These problems mean that further development work and testing would be necessary to ensure that such a system was feasible and able to be used across all clinical systems.

Response rates were the main variable of interest of this study, and it would be desirable for response rates, particularly those at baseline, to be higher. However, other general practice-based surveys such as the GP Patient Survey have achieved a similar response rate [14, 15], suggesting that such a low response rate would be typical. The advantage of this study is that the potential for greater reliability of the diagnosis obtained from GP clinical systems over the self-reported diagnoses in the GP Patient Survey. The response rates for the data collection of the four elective PROMs were higher (60.3%) than the baseline response rates reported here, as were the response rates post-surgery (87.5%) [16]. Nevertheless, non-response bias has been observed in relation to outcomes in elective surgery [17]. Overall response rates to health surveys are decreasing [18], but the magnitude of response rates are not always thought to be a problem. Roland and colleagues found little evidence that low response rates and non-response bias in health surveys led to unfairness in payments for performance in general practices [14]. It is not known whether there is any non-response bias in this study, and efforts to enhance response rates should not be dismissed. Traditional methods to increase response rates include a good cover letter and sending a reminder.

More modern methods, however, such as using electronic surveys may help increase response rates, especially in younger age groups. In this respect, it is worth noting that the response rate at follow-up of 71.5% suggests that once patients are 'on-board' the study they are likely to continue to respond, even one year later.

The other variables of interest in this study were the EQ-5D mean and VAS scores reported by the survey respondents, with very little change occurring in health-related quality of life. Change in the EQ-5D in LTCS has been considered in the Health Survey for England with little change shown [19], however the cross-sectional design of the Health Survey does not allow for assessment of changes within the same cohort. In view of the goal of the current policy in England to assess performance of the health service by means of the EQ-5D, these results raise broader questions about optimal indicators of care for long-term conditions and appropriate timelines for assessment. Currently there are a number of major uncertainties regarding trends and timelines for the progression of LTCs. Moreover, given evidence from this study that better baseline EQ-5D scores resulted in higher response rates at follow-up, it is unclear how helpful self-reported health-related quality of life is in monitoring such trends.

To improve response rates, the invitation to contribute self-reports of health needs to be more engaging in the sense of patients perceiving that their reports will actually serve a purpose. This greater sense of point or purpose could be achieved in three distinct, but not mutually exclusive ways. Firstly, patients could find the information valuable and informative, for example by providing feedback of their progress over time or in comparison with other patients. Secondly, patients could find the information in PROMs helpful in preparing for consultations with healthcare providers or as a part of regular review or assessment. Thirdly, it is conceivable that patients would value providing information if it were truly the case that PROMs data were used to provide evidence of quality or performance of services. The NHS and indeed all other healthcare systems are a long way from being able to support any of these possible uses of PROMs for long-term conditions that might enhance patient engagement. Further research needs to address whether PROMs can better inform patients about their progress, support communication of need or facilitate contributions to quality assessment.

Acknowledgments

We would like to thank the Primary Care Research Networks in London and the North-West; the Diabetes Research Network in London; all the practice staff who helped with the recruitment of practices and patients, and all the patients who participated in the surveys. Furthermore, we would like to thank Mrs Sarah Dummett for her support in collecting the data.

Competing interest declaration

All authors have completed the Unified Competing Interest form at www.icmje.org./coi_disclosure.pdf (available on request from the corresponding author) and declare that (1) All authors have support from the Department of Health (England) for the submitted work; (2) none of the authors have a relationships with commercial companies that might have an interest in the submitted work in the previous 3 years; (3) their spouses, partners, or children have no relationships that may be relevant to the work; and (4) none of the authors have non-financial interests that may be relevant to the submitted work.

Details of contributions

RF, CJ, and HD conceived and designed the study. MP and HC collected data. All authors contributed to the data analysis. MP wrote the article, with contributions from HC, HD, CJ and RF. All authors had access to the data in the study, and can take responsibility for the integrity of the data and the accuracy of the data analysis. All authors edited and approved the final version of the manuscript.

Funding and sponsor

This research was initially funded as an individual project under the Department of Health's Policy Research Programme. From 2011, it was incorporated within the research programme of the Department of Health funded Policy Research Unit on Quality and Outcomes of Person Centred Care (QORU), a collaboration between the London School of Economics and Political Science (LSE) and the Universities of Kent and Oxford.

The University of Oxford Clinical Trials and Research Governance (CTRG) acted as sponsor.

Disclaimer

This is an independent report commissioned and funded by the Policy Research Programme in the Department of Health. The views expressed are not necessarily those of the Department.

Ethics approval

Ethics approval was given by the National Research Ethics (NRES) Committee of the Isle of Wight, Portsmouth & South East Hampshire (now the NRES South-Central Committee) in March 2010 (reference number 10/H0501/10). A change of the diabetes PROM was approved in September 2010.

For the baseline, consent was implied by completion and return of the survey. Consent to the follow-up was given by the majority of baseline participants by providing the research team with contact details as part of completing the baseline questionnaire.

Registration

This study was registered on the National Institute for Health Research (NIHR) portfolio in June 2010 (UKCRN ID: 8462).

Data sharing

No additional patient-level data is available. The full report of the study will be made available on the website of the Policy Research Unit on Quality and Outcomes of Person Centred Care (http://www.qoru.ac.uk/).

Figure legend

Figure 1: Recruitment and participation (n) of people with LTCs

REFERENCES

1. Barnett K, Mercer SW, Norbury M, et al. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012, **380**:37-43

2. Department of Health. Long Term Conditions Compendium of Information. Third ed. London: Department of Health, 2012.

3. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007;**22 Suppl 3**:391-5

4. Pearson WS, Bhat-Schelbert K, Probst JC. Multiple chronic conditions and the aging of America: A challenge for primary care physicians. Journal of Primary Care and Community Health 2012;**3**(1):51-56

5. Anderson G, Horvath J. The growing burden of chronic disease in America. Public Health Rep 2004;**119**(3):263-70

6. Heyworth IT, Hazell ML, Linehan MF, et al. How do common chronic conditions affect health-related quality of life? Br J Gen Pract 2009;**59**(568):e353-8

7. Epping-Jordan JE, Pruitt SD, Bengoa R, et al. Improving the quality of health care for chronic conditions. Qual Saf Health Care 2004;**13**(4):299-305

8. World Health Organization. Innovative care for chronic conditions: building blocks for action. Geneva, 2002.

9. Department of Health. Equity and Excellence: Liberating the NHS. London, 2010.

10. Department of Health. The NHS Outcomes Framework 2012/2013. London, 2011.

11. Black N, Jenkinson C. How can patients' views of their care enhance quality improvement? BMJ 2009;**399**:202-05

12. Department of Health. The NHS Outcomes Framework 2011/2012. London, 2010

13. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001;**33**(5):337-43

14. Roland M, Elliott M, Lyratzopoulos G, et al. Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. BMJ 2009;**339**:b3851

15. Campbell J, Smith P, Nissen S, et al. The GP Patient Survey for use in primary care in the National Health Service in the UK--development and psychometric characteristics. BMC Fam Pract 2009;**10**:57

16. Browne J, Jamieson L, Lewsey J, et al. Patient reported outcome measures (PROMs) in elective surgery. Report to the Department of Health, 2007.

17.Hutchings H, Neuburger J, Grosse Frie K, et al. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. Health and Quality of Life Outcomes 2012, 10:34 doi:10.1186/1477-7525-10-34

18. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007;**17**(9):643-53

19. Department of Health. Setting Levels of Ambition for the NHS Outcomes Framework. A technical annex to support Delivering our NHS care objectives: a consultation on the draft mandate

to the NHS Commissioning Board Chapter 4: Enhancing quality of life for people with long-term conditions. London, 2012

The routine collection of patient-reported outcome measures (PROMs) for long-term conditions in primary care: a cohort survey

Michele Peters [1], Senior Research Scientist
Helen Crocker [1], Researcher
Crispin Jenkinson [1], Professor of Health Services Research
Helen Doll [2], Senior Statistician
Ray Fitzpatrick [1], Professor of Public Health and Primary Care

[1] Nuffield Department of Population Health, University of Oxford, Old Road Campus, Oxford OX3 7LF, UK
[2] Oxford Outcomes, Seacourt Tower, West Way, Oxford OX2 0JJ, UK

Corresponding author

Michele Peters michele.peters@dph.ox.ac.uk

Telephone: 01865 289428

Nuffield Department of Population Health, Health Services Research Unit, University of Oxford, Old Road Campus, Oxford OX3 7LF

Keywords: Patient-reported outcome measures, long-term conditions, cohort study, primary care

Word count: 3327

ABSTRACT

Objectives

To evaluate the feasibility of using patient-reported outcome measures (PROMs) for long-term conditions (LTCs) in primary care

Design

A cohort postal survey conducted from September 2010 to April 2012

Setting

Primary care practices (n=33) in London and the North-West of England

Participants

4484 patients with a diagnosis of asthma, chronic obstructive pulmonary disease (COPD), diabetes, epilepsy, heart failure or stroke were sent a survey at baseline

Main outcome

The main outcome was to evaluate the feasibility of and the recruitment strategies for collecting PROMs data in LTCs by assessing the rResponse rates for the baseline and follow-up surveys.

Secondary outcomes were the evaluation of change scores of, and the EQ-5D index and visual analogue scale (VAS), a generic patient-reported outcome measurebetween baseline and follow-up surveys.

Results

The baseline survey achieved a response rate of 38.4% (n=1721/44854) and at follow-up 71.54% (n=1136/158990). Response rates varied by LTC. Little change was found in health-related quality of life for the total sample (-0.001 for the EQ-5D index score and 0.12 for the EQ-5D Visual Analogue ScaleVAS) between patients responding to both the baseline and follow-up surveys.

Conclusions

The response rate to the baseline survey was similar to that of other general practice surveys. Current UK policy aims to assess health service performance in LTCs by means of using PROMs. It thus would be desirable to improve response rates by making the invitation to self-reports of health-related quality of life more engaging for patients. Results of the studyon the EQ-5D score raise questions about optimal indicators for LTCs and appropriate timelines for assessment.

ARTICLE SUMMARY

Focus

- Long-term conditions (LTCs) reduce people's quality of life and increase the use of health care services
- The NHS Outcomes Framework aims to enhance quality of life in LTCs as assessed by patient-reported outcome measures (PROMs), in particular the EQ-5D

Key messages

- This survey's response rate of 38.4% at baseline was similar to response rates in other general practice surveys
- In those responding, little change was found in health-related quality of life over the course of one year
- In view of current policy in England, the response rates and limited change in health-related quality of life raises questions about the optimal indicators of care and appropriate timelines for assessment

Strengths and limitations

- The main strength of the study is that it is the first empirical evidence on the use of PROMs for LTCs in primary care focusing on patients with a confirmed diagnosis
- Another strength is that the study provides evidence that patients with LTCs who have completed a PROM once, are likely to complete a follow-up
- Limitations include the response rate, some logistic problems with identifying eligible patients from the GP databases and the number of patients excluded by the practices.

INTRODUCTION

The prevalence of long-term or chronic conditions, including multi-morbidity, is increasing world-wide [1-5]. Long-term conditions (LTCs) usually reduce people's quality of life, particularly in the case of multi-morbidity [2, 3, 6]. People with limiting LTCs are the most intensive users of the most expensive health care services [2] and the majority of health expenditure is for people with chronic conditions [5].

Chronic conditions are increasingly becoming the main concern of health care systems [5, 7] and a current challenge is to evaluate models of care and to develop new models that are more cost effective [4]. The Innovative Care for Chronic Conditions (ICCC) Framework aims to address different levels of health care including that at the micro level focused on patients and families.[8] In England, improving health care outcomes for all is the primary purpose of the National Health Service (NHS) in England [9] and the NHS Outcomes Framework [10] sets out the indicators for measuring health outcomes in NHS services. The Outcomes Framework's second domain for improvement is the enhancement of quality of life for people with LTCs as assessed by patient-reported outcome measures (PROMs).

Patients' views are complementary to those of clinicians and provide unique information into the humanity and effectiveness of health care.[11] Routinely collecting data on patients' views through PROMS provides an opportunity to help drive change in the organisation and delivery of health care. Since 2009, PROMs have been used on a routine and mandated basis to assess outcomes in four elective surgical procedures in the NHS in England (http://www.ic.nhs.uk/proms, accessed 21.3.13). The role of PROMs in LTCs is more challenging than in elective surgery. In elective surgery, the PROMs are used to help assess the effectiveness of a single, discrete procedure. By contrast, LTCs are complex to manage involving diverse service providers and interventions over long timelines. The purpose of this study was to assess the feasibility of using PROMs to monitor ~~the~~ aggregated quality of life data ~~of individuals with~~for a range of LTCs in primary care.

METHODS

A cohort survey was conducted in primary care from September 2010 to April 2012, with baseline and follow-up data being collected one year apart. The primary aim of the study was to evaluate the recruitment strategies for and the feasibility of collecting PROMs data in LTCs by assessing response rates at two points of data collection. Secondary outcomes were the change scores of the EQ-5D index and visual analogue scale (VAS) between baseline and follow-up surveys. Two considerations informed the secondary outcomes of the study: on the one hand the NHS Outcomes Framework's [12] second domain refers to the possibility of changes over time in the quality of life in LTCs and on the other hand, the trajectory of many LTCs is deterioration in health status. The LTCs included were asthma, chronic obstructive pulmonary disease (COPD), diabetes, epilepsy, heart failure and stroke. Ethics approval had been obtained through the National Research Ethics (NRES) Committee of the Isle of Wight, Portsmouth & South East Hampshire (now the NRES South-Central Committee) in March 2010 and R&D approval from 20 participating Primary Care Trusts (PCTs). The study was registered on the National Institute for Health Research (NIHR) portfolio (UKCRN ID: 8462).

Formatted: Space Before: 0 pt

Design

A cohort survey was conducted by post. The baseline survey was provided pre-packed to participating practices and was mailed by practice staff, accompanied by a covering letter from the general practitioner. The baseline survey was returned to the Oxford research team and the follow-up survey was sent from Oxford. Consent was implied by returning the baseline survey. If participants were willing to complete a follow-up questionnaire, they were asked to give consent and to provide the Oxford team with their contact details. All surveys were numbered. This served a dual purpose, firstly to identify the respondents' practice and secondly to match baseline and follow-up responses. A 'thank you/reminder' letter was sent by the practices two weeks after the mailing of the baseline survey to all patients invited into the study. The follow-up reminder, also sent after two weeks, was sent by the research team and targeted at non-responders only. A Microsoft Access database was set up to manage the mailing and receipt of the surveys.

Setting

The study was conducted in 33 primary care practices in London (n=18) and the North-West of England (NW) (n=15). Practices were recruited with the support of PCT staff, research nurses, local Primary Care Research Networks and the Diabetes Research Network in London. For less prevalent LTCs a larger number of practices participated. The majority of practices covered 3 LTCs and one practice covered 2 LTCs. Power calculations determined that a total of 1050 patients (or 175 per LTC) were needed to detect, at two-sided $p<0.05$ with 80% power, both a 15% difference in response rates between two independent groups and a moderate effect size of 0.3 in PROM score change over time. Ten practices provided patients for asthma (5 in London and 5 in NW), 16 for COPD (8 in London and 8 in NW), 10 for diabetes (5 in London and 5 in NW), 23 for epilepsy (13 in London and 10 in NW), 20 for heart failure (11 in London and 9 in NW) and 19 for stroke (12 in London and 7 in NW). The number of practices per LTC varied due to the prevalence of the LTC and practice size. Practices varied in size: 12 were small (<5800 patients), 13 medium (5800-10,500 patients) and 8 large (>10,500 patients). A slightly larger number of practices were recruited from more deprived areas (Table 1).

Table 1: Number of practices per social deprivation quintile

	Quintile	Range (IMD rank 2010*)	London	NW	Total
Most deprived	**Q1**	1 – 6496	4	4	8
	Q2	6497 - 12992	4	6	10
	Q3	12993 - 19488	5	1	6
	Q4	19489 - 25984	4	2	6
Least deprived	**Q5**	25985 - 32482	1	2	3

* Indices of Multiple Deprivation 2010

Participants

Eligible patients were identified through an automatic and remote search of practices' clinical systems by a subcontracted IT company prior to the baseline survey. The search is specific to every clinical system of which a range is available. The search was developed only for the 3 most widely used clinical systems. The aim was to identify approximately 50% of the patients for each LTC in every practice by a selection based on odd or even months of patients' birthdays. The search was based on Read codes in line with the criteria used in the Quality and Outcomes Framework (QOF)

with two exceptions: patients with diabetes needed to be 18 years of age and patients with transient ischaemic attack(s) (TIAs) were excluded from the stroke group. The number of patients extracted by the search was compared to QOF estimates to check the expected numbers of patients were being selected. Patients were included in the survey for one LTC only; if they had multiple LTCs they were included for the rarest condition.

A member of staff (usually a GP or a nurse) reviewed the list of eligible patients identified from clinical systems to exclude any patients who were not considered suitable to be invited into the study. The instruction to practices was to exclude patients ~~(particularly~~ if participation in the survey might cause serious distress. The practice staff made the judgement of when this criterion applied~~)~~.

Patient-reported outcome measures (PROMs)

A generic PROM and appropriate disease-specific PROMs were included in the surveys, as well as standard demographics questions and a question on comorbidities. The PROMs were selected on the basis of their psychometric properties evaluated by review work (http://phi.uhce.ox.ac.uk/newpubs.php , accessed 13.12.12) and licensing agreements. This article presents the findings from the EQ-5D [13~~2~~], the generic measure used. The EQ-5D is a measure of health status primarily designed to provide a single-index value which represents the utility of specific health states, i.e. how given health states are valued by the general population.[13~~2~~] It takes approximately five minutes to complete and comprises five items, one each on mobility, self-care, usual activities, pain/discomfort and anxiety/depression, that are all scored on a three-point scale. A single-index value is calculated from the five items typically with a score range from 1 (perfect health) to 0 (death) although a small number of scores below 0 can be obtained indicating states worse than death. The EQ-5D also includes a Visual Analogue Scale (VAS), ranging from 0 'worst imaginable health state' to 100 'best imaginable health state'.

Analysis

Response rates were calculated overall and for each LTC at baseline and follow-up. Cross-tabulations, with chi-square tests, were used to assess differences in response rates at baseline, consent to be followed up and response rate at follow-up between categorical variables (participant LTC, gender, age, location, ethnicity), and t-tests were used to compare mean scores for continuous data (time since diagnosis, number of comorbidities). Analysis of variance was used to identify factors (type of LTC, practice or patient related-factors or mean baseline EQ-5D scores) that were significantly related to follow-up response rates. ~~The level of significance was set at two-sided p<0.05.~~ All the variables relevant to cohort follow-up response rates were entered into a logistic regression (no consent to follow-up=0 vs. consent for follow-up=1, and follow-up not completed=0 vs. follow-up completed=1). The level of significance was set at two-sided p<0.05.

Changes in the EQ-5D utility and VAS scores for respondents to both the baseline and follow-up surveys were assessed with paired t-tests. No data imputation was performed so the number of patients in the analysis of the EQ-5D may be lower than the number of respondents to the survey. The level of significance was set at p<0.05.

RESULTS

Recruitment of participants

The total number of patients extracted from each search was compared to 50% of the total QOF estimate (by LTC) based on the 2009/2010 QOF prevalence rates (http://www.qof.ic.nhs.uk/search/). The number of patients extracted by the searches is higher (total n=17,455), than those invited into the study (n=5596), as the search data was available for all LTCs for every practice but surveys were sent only to up to three LTCs per practice. The flow of the 5596 patients invited into the study is illustrated in Figure 1. The QOF estimates, number of patients extracted from the search, the overall difference and range of difference between the practices are presented in Table 2. A positive score on the difference means that a higher number of patients than expected were extracted from the search, whereas a negative number means that a lower than expected number was extracted. Although these results give an indication of how well the searches worked, they need to be interpreted with caution. The number of patients extracted in small practices was low. The extracted numbers for diabetes and stroke were expected to be lower than the number of patients in QOF. Diabetes patients aged 17 and upwards are included in QOF, whereas this study only included diabetes patients aged 18 or over. This difference is expected to be minimal. Stroke presented more of a challenge as QOF includes TIAs which were not included in this study. Therefore, it was unlikely that the numbers extracted from the search would closely match the QOF estimates.

Table 2: Numbers of eligible patients extracted in PROMs pilot search compared to QOF estimates

LTC	QOF estimate (n)	PROMs pilot search (n)	Difference		Range of difference for practices (%)	
			n	%		
Asthma	7722	7615	-107	-1.4	-48.1	+62.6
COPD	2036	1590	-446	-21.9	-71.2	+42.1
Diabetes	5878	5154	-824	-13.8	-28.6	+19.1
Epilepsy	822	1088	+266	+32.4	-80.0	+348.3
Heart failure	948	937	-11	-1.2	-29.4	+17.7
Stroke	2351	1071	-1280	-54.4	-71.2	-30.0

Exclusions by practices

The rates of patients excluded from being sent a questionnaire varied between LTCs and practices (Table 3). The rate of exclusions tended to be higher in the NW than in London, although this difference was small for asthma, COPD and diabetes. However, the difference was substantial for epilepsy (56.6% in NW vs. 10.0% in London), heart failure (28.5% NW vs. 12.9% London) and stroke (24.2% NW vs. 16.8% London). The reasons for excluding patients were death, blindness, co-morbidities (such as dementia or cerebral palsy) or learning difficulties, patients having moved or being on an extended holiday, not having a clear diagnosis, recently having a traumatic life event or being in a care home. Comorbidities and learning difficulties were particularly common reasons for excluding epilepsy patients.

Table 3: Exclusions of patients by practices (cohort baseline data)

LTC (n practices)	N patients extracted		N patients excluded		% included	
	Total	Practice range	Total	Practice range	Total	Practice range
Asthma (10)	1628	64-684	294	0-197	81.9	63.5-100.0

COPD (16)	602	8-88	35	6-80	94.2	81.5-100.0
Diabetes (10)	1169	63-185	48	0-13	95.9	90.2-100.0
Epilepsy (23)	985	4-260	460	4-78	53.3	19.2-100.0
Heart failure (20)	687	5-143	167	4-81	75.7	56.6-100.0
Stroke (19)	525	4-69	107	0-19	79.6	52.4-100.0
Total (33)	5596	38-880	1111	0-261	80.1	35.5-100.0

Response rates

Cohort baseline

A total of 44856 questionnaires were sent and 1721 were returned achieving an overall response rate of 38.4%. Thirteen patients were excluded from analysis and follow-up as they reported not to have been diagnosed with the specified LTC. The response rate varied between LTCs, with the response rate being the highest in heart failure (50.4%, n=262) and the lowest in asthma (30.0%, n=400) (Table 4). There was significant variation in response rates between LTCs (p<0.001), by practice (p=0.018) and by region (p=0.002). Additionally, across all LTCs there were significant non-linear relationships between response rates and the practices' deprivation score (p=0.024, Table 5), and the practices' EQ-5D mean score adjusted for patient age, gender, time since diagnosis and number of comorbidities (p=0.004). Practices with a QOF score of 100 (maximum score) had significantly (p=0.013) higher response rates (mean response rate=42.9%, SD 11.9) than those who did not (mean=35.5%, SD 14.3).

Table 4: Cohort baseline survey numbers sent and returned, and response rates (%) for each LTC (overall, by region and practice range).

LTC (n practices)	N		Overall RR (%)	Regional RR (%)		RR by practice (range) (%)
	Sent	Returned		London	NW	
Asthma (10)	1334	395400	30.0	22.7	33.0	14.3-50.0
COPD (16)	567	279	49.2	43.0	54.3	32.1-66.7
Diabetes (10)	1121	448	40.0	30.5	50.1	28.1-61.3
Epilepsy (23)	525	180	34.0	35.5	33.6	0-53.9
Heart failure (20)	520	262	50.0	48.8	51.1	30.6-71.4
Stroke (19)	418	152	36.4	30.0	44.0	7.7-63.2
TOTAL (33)	4485	1721	38.4	32.7	42.2	14.2-58.9

Table 5: Mean response rate by deprivation score (N refers to the number of practices for each LTC covered i.e. 987 cases)

Deprivation quintile	N	Mean	SD
Most deprived	23	34.2	13.5
2	2930	43.6	13.5
3	18	39.7	10.2
4	18	43.2	9.8
Least deprived	9	48.0	14.1

Cohort follow-up

At baseline, ~~93.1~~92.3% of responding patients (~~1603~~ 1589 of 1721) had agreed and were eligible to be sent a follow-up questionnaire. The lowest rate of consent was achieved for heart failure (90.5%) and the highest for COPD (95.3%). ~~Thirteen patients were excluded from follow-up as they indicated that they had not been diagnosed with the LTC referred to; the denominator for the follow-up response rate was thus 1590.~~ For the total sample, there was no significant difference in whether consent to follow-up was given by LTC, age, region, ethnicity, number of comorbidities or time since diagnosis. There was, however, a difference between practices (p=0.008) with the proportion of baseline respondents giving consent ranging from 78.6% to 100%. There were some ~~disease-specific~~ differences in consent for individual LTCs, including gender (96.1% of men consented vs. 90.1% of women, p=0.029) and practice (consent in asthma ranged from 79.1%-100% in 10 practices, p=0.010) in asthma, and number of comorbidities for epilepsy (mean 1.17, SD 1.30 for those who consented vs. mean 2.14, SD 2.00 for those who did not consent, p=0.017) and heart failure (mean 1.92, SD 1.62 for those who consented vs. mean 2.60, SD 1.71 for those who did not consent, p=0.049) respectively.

The overall response rate of baseline respondents who consented to the follow-up was 71.54% (n=1136/158~~990~~). Numbers of questionnaires sent and response rates by LTC are given in Table 6. Differences in response rates to the follow-up survey were examined by means of baseline characteristics of respondents who had consented to the follow-up. For the overall sample, there were significant differences in response rate at follow-up by LTC (p=0.015), age (p<0.001), ethnicity (p=0.008) and region (p=0.007), with epilepsy and heart failure patients being less likely to respond than patients with one of the other LTCs, as were younger patients, those based in London and those from ethnic minority backgrounds (Table 7). The baseline mean EQ-5D score was significantly lower (p<0.001) in non-responders to follow-up (mean 0.66, SD 0.33) than in follow-up responders (mean 0.73, SD 0.29). There were no significant differences in the response rate at follow-up for gender, time since diagnosis, number of comorbidities or practice.

Table 6: Cohort follow-up survey numbers sent and returned, and response rates (%) for each LTC (overall, by region and practice range).

LTC (n practices)	N		Overall RR (%)	Regional RR (%)		RR by practice (range) (%)
	Sent	Returned		London	NW	
Asthma (10)	366	267	73.0	65.4	75.1	53.9-82.4
COPD (16)	262	187	71.4	68.9	73.1	44.4-82.9
Diabetes (10)	424	321	75.7	67.9	80.7	60.4-87.8
Epilepsy (23)	166	104	62.7	69.8	58.3	25.0-100.0
Heart failure (20)	234	155	66.2	60.6	68.7	25.0-100.0
Stroke (19)	137	102	74.5	70.5	77.6	0-100.0
Total (33)	1589	1136	71.4	67.3	73.7	53.3-66.2

Table 7: Factors significantly related to the number of questionnaires returned at cohort follow-up

		% responders
LTC (p=0.015)	Asthma	73.0
	COPD	71.4
	Diabetes	75.7
	Epilepsy	62.7
	Heart failure	66.2

	Stroke	74.5
Age (years) (p<0.001)	18-24	37.5
	25-34	48.4
	35-44	63.4
	45-54	70.1
	55-64	78.8
	75-84	71.7
	85+	65.2
Region (p=0.007)	London	67.3
	North-West	73.7
Ethnicity (p=0.008)	White	72.9
	Other	63.0

When analyses were performed of prediction of return of the follow-up questionnaire for each of the six LTCs, significant differences were observed for some factors, including age (p<0.001) and ethnicity (p=0.009) in asthma; age (p=0.012) in COPD, gender (p=0.032) and region (p=0.003) in diabetes; age (p=0.011) in epilepsy; and ethnicity (p=0.003) in stroke. The direction of these differences was the same as for the overall sample (i.e. lower completion rate in younger patients and ethnic minorities), with the additional finding that in diabetes, women were less likely to complete the follow-up than men. When entered into a logistic regression (follow-up not completed =0 vs. follow-up completed =1), only some age groups and EQ5D remained significant (**Table 8**). Asthma and age 18-24 served as reference categories.

Table 8: Factors related to completion of the follow-up questionnaire

		p	Odds ratio
LTC	**Asthma**	NS	--
	COPD	NS	1.07
	Diabetes	NS	0.80
	Epilepsy	NS	1.06
	Heart Failure	NS	0.75
	Stroke	NS	0.66
Age (years)	**18-24**	<0.001	--
	25-34	0.001	0.24
	35-44	0.005	0.37
	45-54	NS	0.71
	55-64	NS	1.04
	65-74	NS	1.68
	75-84	0.035	1.70
	85+	NS	1.23
	Region	NS	1.20
	EQ5D	<0.001	2.11
	Ethnicity	NS	0.77
	Gender	NS	0.88
	Constant	NS	1.32

EQ-5D scores

EQ-5D scores could be calculated for 93.1% (n=1058) of the respondents to both the baseline and follow-up surveys. The EQ-5D scores and VAS were found to be significantly different between LTCs both at baseline and follow-up (all p<0.001). However, no significant differences were found for the mean EQ-5D scores for the total sample between baseline and follow-up. For the VAS, a significant difference was found for heart failure only, with scores deteriorating at follow-up (mean difference -3.53, 95% CI -6.67 to -0.38). The data are presented in Table ~~8~~9.

Table ~~8~~9: EQ-5D and VAS mean, and change, scores at baseline and follow-up

		Asthma	COPD	Diabetes	Epilepsy	Heart failure	Stroke	Total
EQ-5D York tariff								
n		255	177	301	95	137	93	1058
Baseline	**Mean**	0.83	0.67	0.73	0.76	0.64	0.67	0.73
	95% CI	0.80-0.86	0.63-0.71	0.69-0.76	0.71-0.82	0.59-0.69	0.61-0.74	0.71-0.75
Follow-up	**Mean**	0.84	0.67	0.72	0.76	0.64	0.67	0.73
	95% CI	0.81-0.87	0.63-0.71	0.69-0.76	0.71-0.81	0.59-0.69	0.60-0.73	0.71-0.75
Difference	**Mean**	0.01	-0.002	-0.003	-0.001	-0.005	-0.008	-0.001
	95% CI	-0.004-0.015	-0.019-0.014	-0.013-0.008	-0.021-0.018	-0.022-0.013	-0.31-0.13	-0.007-0.045
p		NS	NS	NS	NS	NS	NS	NS
	EQ-5D VAS							
n		248	173	296	91	145	82	1035
Baseline	**Mean**	73.77	62.29	68.16	71.40	62.20	73.84	68.54
	95% CI	71.31-76.23	59.31-65.27	65.84-70.48	67.11-75.68	58.93-65.47	67.18-76.75	67.27-69.81
Follow-up	**Mean**	74.33	62.14	69.76	73.59	58.67	71.96	68.42
	95% CI	71.94-76.72	59.13-65.16	67.53-71.99	69.68-77.50	55.10-62.23	67.18-76.75	67.16-69.68
Difference	**Mean**	0.56	-0.15	1.60	2.20	-3.53	-1.88	0.12
	95% CI	-1.09-2.21	-2.94-2.66	-0.19-3.39	-0.96-5.36	-6.67--0.38	-5.12-1.37	-0.87-1.10
p		NS	NS	NS	NS	0.029	NS	NS

DISCUSSION

This study provides the first empirical evidence on the use of PROMs for LTCs in primary care focusing on patients with a confirmed diagnosis. The evidence on participation of individuals with LTCs is essential for making decisions about the feasibility of expanding the current PROMs programme to LTCs. The aim of the study was to evaluate the feasibility of collecting PROMs data in LTCs through primary care focusing on response rates to a baseline and follow-up survey conducted one year apart. Overall, a 38.4% response rate was achieved at baseline and 71.5% for the cohort (those responding at baseline and agreeing to be sent a follow-up questionnaire) at follow-up. Response rates varied between LTCs at both assessments. Other factors significantly related to response rates were practice factors (including deprivation score, QOF score and mean EQ-5D score) at baseline, and respondent factors (including age, ethnicity, region and mean baseline EQ-5D score) at follow-up. Other methods to assess feasibility included testing the remote and automatic search of GP databases for eligible patients and to monitor patient exclusions. Problems were encountered with the search. It was possible to overcome these to some extent by the IT company amending the

searches. However, discrepancies remained in the number of patients extracted from the search when compared with QOF estimates. The PROMs programme aims to be as inclusive as possible and therefore the number of patients excluded by practices at baseline was monitored. The exclusion rates varied by LTC and by region. The variation by LTC may be related to the nature of the specific LTC or the comorbidities associated with a particular LTC. The highest exclusion rate was for epilepsy, which practices reported to be expected as many people with epilepsy have learning difficulties or other comorbidities which would make it inappropriate for them to be included in the study. There were no significant changes in EQ-5D scores between baseline and follow-up. The VAS was significantly different between baseline and follow-up for heart failure only.

The main strength of this study is the potential for greater reliability of the diagnosis obtained from GP clinical systems over the self-reported diagnoses in the GP Patient Survey [143, 154]. While the study was able to detect patients with LTC using remote and automatic search of GP databases, weaknesses of the methods tested in this study were the lack of information about non-responders of the baseline survey, high exclusion rates by practices in some LTCs and the logistics of remotely identifying eligible patients from GP databases. No information about non-responders had been collected as there were concerns that collecting these data was too burdensome for participating practices. However, such data would have been valuable to assess non-response bias beyond the non-response bias from the follow-up. Exclusions rates varied between practices. This was partly due to problems with identifying eligible patients in the practices where the search by the IT company was first conducted. For example, in the case of epilepsy the initial search did not include the code in relation to medication (as by QOF) and some practices chose for a member of staff to exclude ineligible patients rather than the search being re-run. Some differences in exclusions occurred due to practices' different interpretations of the instruction to exclude patients in whom invitation into the study would cause serious distress, although this is believed to only have had a minor impact. Reasons for exclusions were not recorded systematically (again to minimize the burden on practices). Practices informally reported that the most ~~C~~common~~ly reported~~ reasons for exclusion were comorbidities such as learning difficulties or mental health problems. This suggests that it may be not be feasible to collect PROMs data from all patients with some diagnoses; other methods, such as collecting data by proxy, may have to be considered. Remotely identifying eligible patients was challenging due to mistakes made in the initial searches (search criteria for asthma and epilepsy did not meet QOF criteria) and the search working less well in some of the clinical systems used by the participating practices. ~~Furthermore, T~~these problems ~~found in identifying eligible patients from the practice databases~~ mean~~s~~ that further development work and testing would be necessary to ensure that such a system was feasible and able to be used across all clinical systems.

Response rates were the main variable of interest of this study, and it would be desirable for response rates, particularly those at baseline, to be higher. However, other general practice-based surveys such as the GP Patient Survey have achieved a similar response rate [134, 145], suggesting that such a low response rate would be typical. The advantage of this study is that the potential for greater reliability of the diagnosis obtained from GP clinical systems over the self-reported diagnoses in the GP Patient Survey. The response rates for the data collection of the four elective PROMs were higher (60.3%) than the baseline response rates reported here, as were the response rates post-surgery (87.5%) [156]. Nevertheless, non-response bias has been observed in relation to outcomes in elective surgery ~~{Hutchings, 2012 #93}~~[17]. Overall response rates to health surveys are decreasing [186], but the magnitude of response rates are not always thought to be a problem. Roland and

colleagues found little evidence that low response rates and non-response bias in health surveys led to unfairness in payments for performance in general practices [134]. It is not known whether there is any non-response bias in this study, and efforts to enhance response rates should not be dismissed. Traditional methods to increase response rates include a good cover letter and sending a reminder. More modern methods, however, such as using electronic surveys may help increase response rates, especially in younger age groups. In this respect, it is worth noting that the response rate at follow-up of 71.5% suggests that once patients are 'on-board' the study they are likely to continue to respond, even one year later.

The other variables of interest in this study were the EQ-5D mean and VAS scores reported by the survey respondents, with very little change occurring in health-related quality of life. Change in the EQ-5D in LTCS has been considered in the Health Survey for England with little change shown [19], however the cross-sectional design of the Health Survey does not allow for assessment of changes within the same cohort. In view of the goal of the current policy in England to assess performance of the health service by means of the EQ-5D, these results raise broader questions about optimal indicators of care for long-term conditions and appropriate timelines for assessment. Currently there are a number of major uncertainties regarding trends and timelines for the progression of LTCs. Moreover, given evidence from this study that better baseline EQ-5D scores resulted in higher response rates at follow-up, it is unclear how helpful self-reported health-related quality of life is in monitoring such trends.

To improve response rates, the invitation to contribute self-reports of health needs to be more engaging in the sense of patients perceiving that their reports will actually serve a purpose. This greater sense of point or purpose could be achieved in three distinct, but not mutually exclusive ways. Firstly, patients could find the information valuable and informative, for example by providing feedback of their progress over time or in comparison with other patients. Secondly, patients could find the information in PROMs helpful in preparing for consultations with healthcare providers or as a part of regular review or assessment. Thirdly, it is conceivable that patients would value providing information if it were truly the case that PROMs data were used to provide evidence of quality or performance of services. The NHS and indeed all other healthcare systems are a long way from being able to support any of these possible uses of PROMs for long-term conditions that might enhance patient engagement. Further research needs to address whether PROMs can better inform patients about their progress, support communication of need or facilitate contributions to quality assessment.

Acknowledgments

We would like to thank the Primary Care Research Networks in London and the North-West; the Diabetes Research Network in London; all the practice staff who helped with the recruitment of practices and patients, and all the patients who participated in the surveys. Furthermore, we would like to thank Mrs Sarah Dummett for her support in collecting the data.

Competing interest declaration

All authors have completed the Unified Competing Interest form at www.icmje.org./coi_disclosure.pdf (available on request from the corresponding author) and declare that (1) All authors have support from the Department of Health (England) for the submitted work; (2) none of the authors have a relationships with commercial companies that might have an interest in the submitted work in the previous 3 years; (3) their spouses, partners, or children have no relationships that may be relevant to the work; and (4) none of the authors have non-financial interests that may be relevant to the submitted work.

Details of contributions

RF, CJ, and HD conceived and designed the study. MP and HC collected data. All authors contributed to the data analysis. MP wrote the article, with contributions from HC, HD, CJ and RF. All authors had access to the data in the study, and can take responsibility for the integrity of the data and the accuracy of the data analysis. All authors edited and approved the final version of the manuscript.

Funding and sponsor

This research was initially funded as an individual project under the Department of Health's Policy Research Programme. From 2011, it was incorporated within the research programme of the Department of Health funded Policy Research Unit on Quality and Outcomes of Person Centred Care (QORU), a collaboration between the London School of Economics and Political Science (LSE) and the Universities of Kent and Oxford.

The University of Oxford Clinical Trials and Research Governance (CTRG) acted as sponsor.

Disclaimer

This is an independent report commissioned and funded by the Policy Research Programme in the Department of Health. The views expressed are not necessarily those of the Department.

Ethics approval

Ethics approval was given by the National Research Ethics (NRES) Committee of the Isle of Wight, Portsmouth & South East Hampshire (now the NRES South-Central Committee) in March 2010 (reference number 10/H0501/10). A change of the diabetes PROM was approved in September 2010.

For the baseline, consent was implied by completion and return of the survey. Consent to the follow-up was given by the majority of baseline participants by providing the research team with contact details as part of completing the baseline questionnaire.

Registration

This study was registered on the National Institute for Health Research (NIHR) portfolio in June 2010 (UKCRN ID: 8462).

REFERENCES

1. Barnett K, Mercer SW, Norbury M, Watt G, Wyke S, Guthrie B. Epidemiology of multimorbidity and implications for health care, research, and medical education: a cross-sectional study. Lancet 2012, **380**:37-43

2. Department of Health. Long Term Conditions Compendium of Information. Third ed. London: Department of Health, 2012.

3. Vogeli C, Shields AE, Lee TA, et al. Multiple chronic conditions: prevalence, health consequences, and implications for quality, care management, and costs. J Gen Intern Med 2007;**22 Suppl 3**:391-5

4. Pearson WS, Bhat-Schelbert K, Probst JC. Multiple chronic conditions and the aging of America: A challenge for primary care physicians. Journal of Primary Care and Community Health 2012;**3**(1):51-56

5. Anderson G, Horvath J. The growing burden of chronic disease in America. Public Health Rep 2004;**119**(3):263-70

6. Heyworth IT, Hazell ML, Linehan MF, Frank TL. How do common chronic conditions affect health-related quality of life? Br J Gen Pract 2009;**59**(568):e353-8

7. Epping-Jordan JE, Pruitt SD, Bengoa R, Wagner EH. Improving the quality of health care for chronic conditions. Qual Saf Health Care 2004;**13**(4):299-305

8. World Health Organization. Innovative care for chronic conditions: building blocks for action. Geneva, 2002.

9. Department of Health. Equity and Excellence: Liberating the NHS. London, 2010.

10. Department of Health. The NHS Outcomes Framework 2012/2013. London, 2011.

11. Black N, Jenkinson C. How can patients' views of their care enhance quality improvement? BMJ 2009;**399**:202-05

12. Department of Health. The NHS Outcomes Framework 2011/2012. London, 2010

13. Rabin R, de Charro F. EQ-5D: a measure of health status from the EuroQol Group. Ann Med 2001;**33**(5):337-43

14. Roland M, Elliott M, Lyratzopoulos G, et al. Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. BMJ 2009;**339**:b3851

15. Campbell J, Smith P, Nissen S, Bower P, Elliott M, Roland M. The GP Patient Survey for use in primary care in the National Health Service in the UK--development and psychometric characteristics. BMC Fam Pract 2009;**10**:57

16. Browne J, Jamieson L, Lewsey J, et al. Patient reported outcome measures (PROMs) in elective surgery. Report to the Department of Health, 2007.

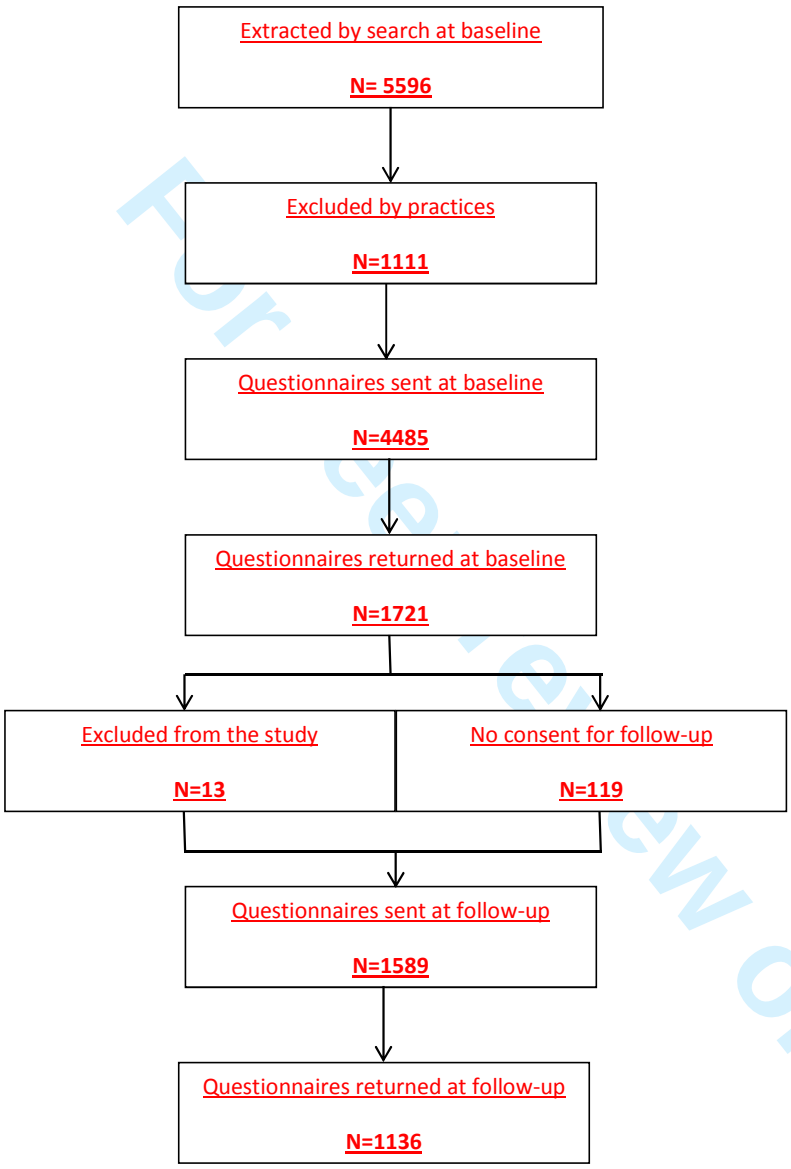
17. Hutchings H, Neuburger J, Grosse Frie K, Black n and van der Meulen J. Factors associated with non-response in routine use of patient reported outcome measures after elective surgery in England. Health and Quality of Life Outcomes 2012, 10:34 doi:10.1186/1477-7525-10-34

18. Galea S, Tracy M. Participation rates in epidemiologic studies. Ann Epidemiol 2007;**17**(9):643-53

19. Department of Health. Setting Levels of Ambition for the NHS Outcomes Framework. A technical annex to support Delivering our NHS care objectives: a consultation on the draft mandate

to the NHS Commissioning Board Chapter 4: Enhancing quality of life for people with long-term conditions. London, 2012

Figure 1: Recruitment and participation (n) of people with LTCs



STROBE Statement—checklist of items that should be included in reports of observational studies

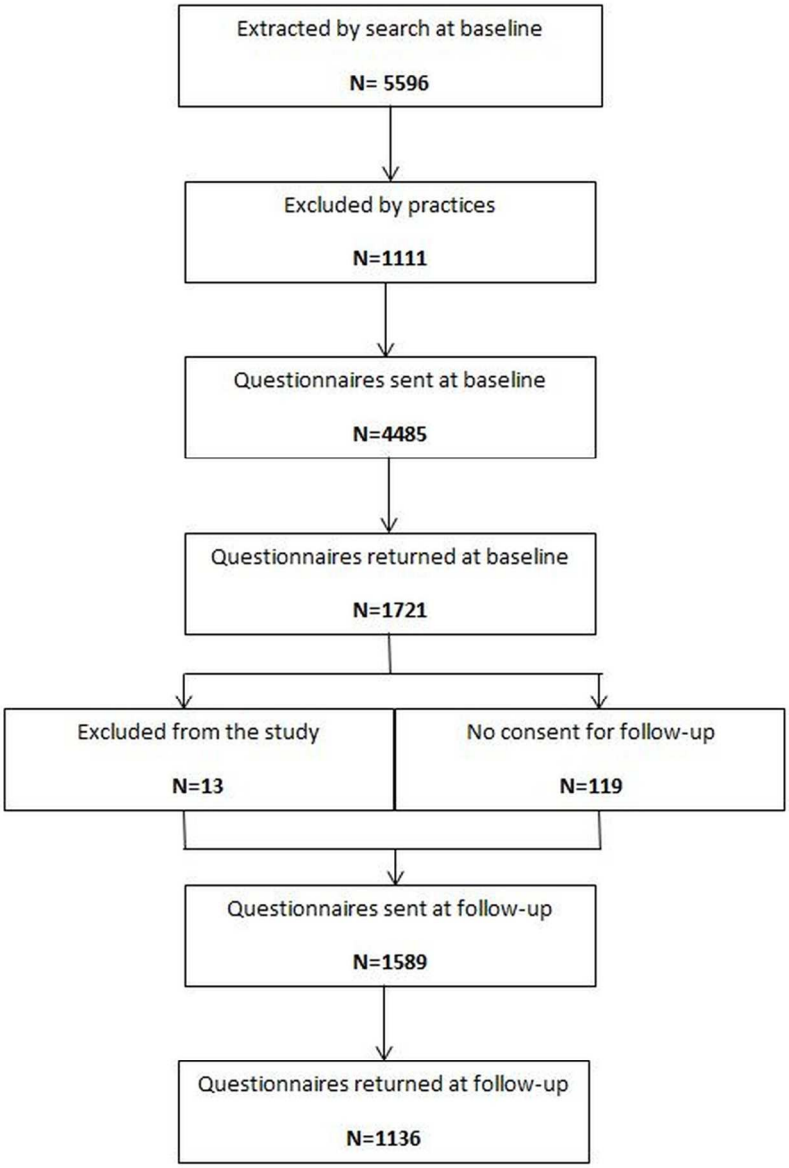
	Item No	Recommendation	
Title and abstract	1	(*a*) Indicate the study's design with a commonly used term in the title or the abstract	✓
		(*b*) Provide in the abstract an informative and balanced summary of what was done and what was found	✓
Introduction			
Background/rationale	2	Explain the scientific background and rationale for the investigation being reported	✓
Objectives	3	State specific objectives, including any prespecified hypotheses	✓
Methods			
Study design	4	Present key elements of study design early in the paper	✓
Setting	5	Describe the setting, locations, and relevant dates, including periods of recruitment, exposure, follow-up, and data collection	✓
Participants	6	(*a*) *Cohort study*—Give the eligibility criteria, and the sources and methods of selection of participants. Describe methods of follow-up *Case-control study*—Give the eligibility criteria, and the sources and methods of case ascertainment and control selection. Give the rationale for the choice of cases and controls *Cross-sectional study*—Give the eligibility criteria, and the sources and methods of selection of participants	✓
		(*b*) *Cohort study*—For matched studies, give matching criteria and number of exposed and unexposed *Case-control study*—For matched studies, give matching criteria and the number of controls per case	N/A
Variables	7	Clearly define all outcomes, exposures, predictors, potential confounders, and effect modifiers. Give diagnostic criteria, if applicable	✓
Data sources/ measurement	8*	For each variable of interest, give sources of data and details of methods of assessment (measurement). Describe comparability of assessment methods if there is more than one group	✓
Bias	9	Describe any efforts to address potential sources of bias	N/A
Study size	10	Explain how the study size was arrived at	✓
Quantitative variables	11	Explain how quantitative variables were handled in the analyses. If applicable, describe which groupings were chosen and why	✓
Statistical methods	12	(*a*) Describe all statistical methods, including those used to control for confounding	✓
		(*b*) Describe any methods used to examine subgroups and interactions	✓
		(*c*) Explain how missing data were addressed	✓
		(*d*) *Cohort study*—If applicable, explain how loss to follow-up was addressed *Case-control study*—If applicable, explain how matching of cases and controls was addressed *Cross-sectional study*—If applicable, describe analytical methods taking account of sampling strategy	✓
		(*e*) Describe any sensitivity analyses	

Continued on next page

Results			
Participants	13*	(a) Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed	✓
		(b) Give reasons for non-participation at each stage	✓
		(c) Consider use of a flow diagram	
Descriptive data	14*	(a) Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders	✓
		(b) Indicate number of participants with missing data for each variable of interest	✓
		(c) *Cohort study*—Summarise follow-up time (eg, average and total amount)	
Outcome data	15*	*Cohort study*—Report numbers of outcome events or summary measures over time	✓
		Case-control study—Report numbers in each exposure category, or summary measures of exposure	
		Cross-sectional study—Report numbers of outcome events or summary measures	
Main results	16	(*a*) Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included	✓
		(*b*) Report category boundaries when continuous variables were categorized	
		(*c*) If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period	
Other analyses	17	Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses	✓
Discussion			
Key results	18	Summarise key results with reference to study objectives	✓
Limitations	19	Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias	✓
Interpretation	20	Give a cautious overall interpretation of results considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence	✓
Generalisability	21	Discuss the generalisability (external validity) of the study results	N/A
Other information			
Funding	22	Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based	✓

*Give information separately for cases and controls in case-control studies and, if applicable, for exposed and unexposed groups in cohort and cross-sectional studies.

Note: An Explanation and Elaboration article discusses each checklist item and gives methodological background and published examples of transparent reporting. The STROBE checklist is best used in conjunction with this article (freely available on the Web sites of PLoS Medicine at http://www.plosmedicine.org/, Annals of Internal Medicine at http://www.annals.org/, and Epidemiology at http://www.epidem.com/). Information on the STROBE Initiative is available at www.strobe-statement.org.



90x124mm (300 x 300 DPI)