

Primary Care Outcomes Questionnaire:

psychometric testing of a new instrument

Abstract

Background

Patients attend primary care for many reasons and to achieve a range of possible outcomes. There is currently no Patient Reported Outcome Measure (PROM) designed to capture these diverse outcomes, and trials of interventions in primary care may thus fail to detect beneficial effects.

Aim

This study describes the psychometric testing of the Primary Care Outcomes Questionnaire (PCOQ), which was designed to capture a broad range of outcomes relevant to primary care.

Design and setting

Questionnaires were administered in primary care in South West England.

Method

Patients completed the PCOQ in GP waiting rooms before a consultation, and a second questionnaire, including the PCOQ and seven comparator PROMs, after 1 week. Psychometric testing included exploratory factor analysis on the PCOQ, internal consistency, correlation coefficients between domain scores and comparator measures, and repeated measures effect sizes indicating change across 1 week.

Results

In total, 602 patients completed the PCOQ at baseline, and 264 (44%) returned the follow-up questionnaire. Exploratory factor analysis suggested four dimensions underlying the PCOQ items: health and wellbeing, health knowledge and self-care, confidence in health provision, and confidence in health plan. Each dimension was internally consistent and correlated as expected with comparator PROMs, providing evidence of construct validity. Patients reporting an improvement in their main problem exhibited small to moderate improvements in relevant domain scores on the PCOQ.

Conclusion

The PCOQ was acceptable, feasible, showed strong psychometric properties, and was responsive to change. It is a promising new tool for assessment of outcomes of primary care interventions from a patient perspective.

Keywords

health services research; patient-centred care; patient-reported outcomes; primary care; psychometrics; quality of care; questionnaires.

INTRODUCTION

Primary care has been evolving in recent years to meet changing population needs and public expectations.¹⁻⁶ As health services globally contend with ageing populations and increasing multimorbidity,⁷ there have been sustained endeavours to improve service quality, costs, and outcomes in primary care. Innovations include electronic consultations,⁸ health coaching, and behavioural change therapies,⁹ and interventions addressing frequent attenders.¹⁰

Assessing the effectiveness of such interventions from a patient perspective involves Patient Reported Outcome Measures (PROMs). Many PROMs are disease specific and tailored to the symptoms and impacts on function of a particular condition.¹¹ These are of limited value in studies where patients have various conditions. As a first-contact, comprehensive, and coordinating service,¹² primary care requires a generic PROM that can be administered regardless of condition. This should be suitable for large-scale trials, based on outcomes that matter to patients and are influenced by GPs. It should also be responsive, that is, able to detect changes over time.¹³

A problem with many generic PROMs is that they are limited to symptoms and function. Primary care patients frequently present with other problems¹⁴ and many have long-term conditions,^{7,15,16} whereby

improvement in function may be unrealistic. Therefore, leading generic PROMs, such as the SF-36 (Short Form-36 Health Survey)¹⁷ and EQ-5D (European Quality of Life-5 Dimensions),¹⁸ often show no change following interventions in primary care.¹⁹⁻²¹ Because of this, alternative measures have been designed specifically for primary care.²¹⁻²³ For example, the Measure Yourself Medical Outcome Profile (MYMOP) allows patients to specify their problems and shows change when other PROMs do not.²¹ However, MYMOP is administered through interviews, making it unfeasible for trials. It also remains limited to symptoms and function. In contrast, the Patient Enablement Instrument (PEI) encompasses broader outcomes related to coping, understanding, and confidence in health, but does not capture symptoms or function. Though it has been validated for primary care,²² the PEI measures outcomes following a single consultation. For many patients, outcomes will become apparent only after longer episodes of care.²⁴ Such outcomes may be multi-layered, capturing aspects of enablement, resilience, symptoms and function, and health perceptions. Without a generic PROM that captures such domains it is impossible to properly assess the outcome of new primary care service configurations from a patient's perspective.

The Primary Care Outcomes Questionnaire (PCOQ) was developed in this context, through a rigorous process²⁵ underpinned

M Murphy, PhD, senior research associate; **S Hollinghurst**, PhD, senior lecturer in health economics; **S Cowlshaw**, PhD, research fellow in primary care; **C Salisbury**, MSc, MD, FRCGP, professor of primary care, Centre for Academic Primary Care (CAPC), Population Health Sciences, Bristol Medical School, University of Bristol, Bristol.

Address for correspondence

Mairead Murphy, Centre for Academic Primary Care, Population Health Sciences, Bristol Medical

School, University of Bristol, Canynge Hall, 39 Whatley Road, Bristol BS8 2PS, UK.

Email: mairead.murphy@bristol.ac.uk

Submitted: 5 May 2017; **Editor's response:** 3 August 2017 **final acceptance:** 8 September 2017.

©British Journal of General Practice

This is the full-length article (published online 27 Mar 2018) of an abridged version published in print. Cite this version as: **Br J Gen Pract 2018;** DOI: <https://doi.org/10.3399/bjgp18X695765>

How this fits in

Patients attend primary care with many types of problems and to achieve a range of possible outcomes, but there is currently no Patient Reported Outcome Measure (PROM) designed to capture these diverse outcomes. As such, trials of interventions in primary care may fail to detect beneficial effects. The Primary Care Outcomes Questionnaire (PCOQ) was developed to measure a range of outcomes commonly influenced in primary care, including health and wellbeing, health knowledge and self-care, confidence in health provision, and confidence in health plan. Testing showed it was acceptable, feasible, and had strong psychometric properties, including responsiveness to change. It is a promising new tool for assessment of outcomes of primary care interventions from a patient perspective.

by a conceptual model of outcomes that included patient health status and ability to impact health status (Figure 1). The authors first interviewed patients and clinicians to establish outcomes that both groups sought to achieve within this framework.²⁶ They then consulted with patients, clinicians, and academics in a Delphi study to identify outcomes most relevant to health and able to be addressed in primary care.²⁷ The authors then developed and tested, through cognitive interviews, a PROM that addressed these outcomes.²⁸ This included health status outcomes, internal features of health empowerment (for example, understanding and ability to self-care), external features (including having access to support and availability of good health care), and outcomes about patients' health perceptions (for example, health concerns and confidence that they are dealing with their health conditions).

The current study reports on the final stage of this process, which aimed to establish the psychometric properties of the PCOQ in primary care patients.

METHOD

Sample and procedures

Waiting room recruitment was chosen as an appropriate method for recruiting patients seeking primary health care. Adult patients were approached in waiting rooms, prior to consultations, in five practices in South West England, and those consulting a GP for themselves were invited to participate. These included a mix of urban and rural, and affluent and deprived areas, with patients from different ethnic backgrounds. Participants self-completed the PCOQ and questions about patient characteristics as they waited for a consultation. Because the baseline questionnaire needed to be short enough to be completed in waiting rooms, the authors collected comparator questionnaires at follow-up only. Patients were asked to take home a second copy of the PCOQ and comparator PROMs for completion 1 week later and return via post. A follow-up email or text reminder was sent after 5 days.

Measures

Prior to testing and item reduction, the PCOQ contained 27 items scored on a 5-point unipolar adjectival scale (from 'No problems' to 'Extreme problems'). The scale wording varies according to attribute, as determined by the qualitative study.²⁶ For example, the item 'How much are you currently affected by pain or discomfort' is anchored at 'Not at all' and 'Extremely'. In contrast, the item 'How much do you understand your health problems' is anchored at 'I understand as much as I want' and 'I understand very much less than I want'. (The PCOQ, and associated scoring files and instructions, are available from the University of Bristol Centre for Academic Care website: <http://www.bristol.ac.uk/primaryhealthcare/resources/pcoq/>.) Patient characteristics were collected at baseline, and seven comparator measures at follow-up. These were: the EQ-5D-5L,²⁹ the Patient Activation Measure (PAM),³⁰ a single item on likelihood of recommending a GP,³¹ the last appointment score,³¹ a single item on support for long-term conditions,³¹ the Illness Perception Questionnaire (IPQ),³² and a single item on change in main problem.³³

Analysis

Psychometric testing of a multi-item

Figure 1. Conceptual models of outcomes influenced by primary care. The darker shaded boxes indicate the constructs that the PCOQ was designed to capture. PCOQ = Primary Care Outcomes Questionnaire.

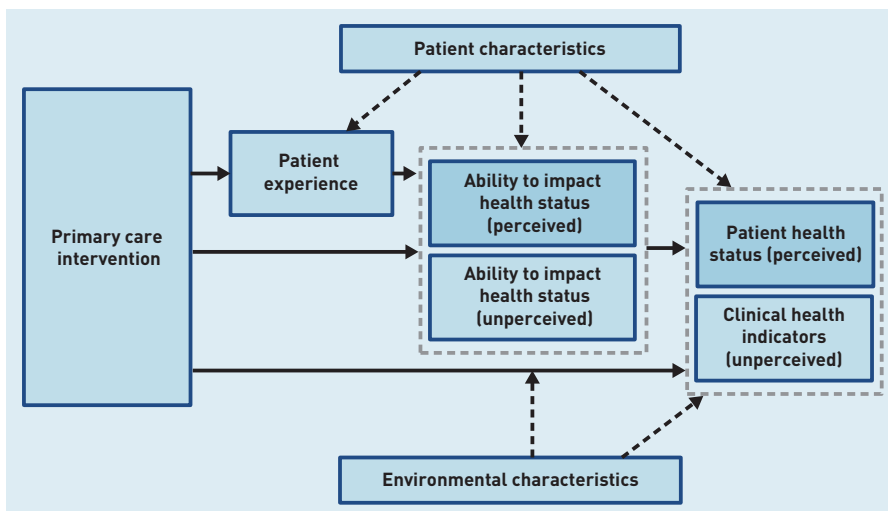


Table 1. Patient characteristics at baseline and follow-up

	Baseline, n (%)	Follow-up, n (%)	Response rate, %	χ^2 (degrees of freedom)	P-value
Total	602	264	44		
Sex				$\chi^2(1) = 0.3$	0.59
Male	233 (39)	107 (41)	46		
Female	366 (61)	157 (59)	43		
Not specified	3 (0)	0 (0)	0		
Age, years				$\chi^2(6) = 25.6$	<0.001
<25	76 (13)	19 (7)	25		
25–34	97 (16)	29 (11)	30		
35–44	98 (16)	34 (13)	35		
45–54	88 (15)	44 (17)	50		
55–64	81 (13)	47 (18)	58		
65–74	81 (13)	51 (19)	63		
≥75	67 (11)	38 (14)	57		
Not specified	14 (2)	2 (1)	14		
Ethnicity				$\chi^2(2) = 0.52$	0.77
White British	498 (83)	226 (86)	45		
White other	38 (6)	15 (6)	39		
Other ethnic group ^a	57 (9)	23 (9)	40		
Not specified	9 (1)	0 (0)	0		
Long-term conditions				$\chi^2(2) = 1.95$	0.38
>1	175 (29)	87 (33)	5		
1	220 (37)	92 (35)	42		
None	200 (33)	82 (31)	41		
Not specified	7 (1)	3 (1)	43		

^aOther ethnic groups included 46% black/African, 25% Asian, 9% mixed race, 16% not specified, 5% other groups.

Box 1. Four dimensions underlying Primary Care Outcomes Questionnaire items

- Factor 1: Health and wellbeing:** Measures overall health status, including symptoms, effects of symptoms on life, and health concerns (8 items).
- Factor 2: Health knowledge and self-care:** Measures health knowledge and patients' ability to self-care and manage symptoms (4 items).
- Factor 3: Confidence in health provision:** Measures patients' confidence in their healthcare providers and ability to access good health care (6 items).
- Factor 4: Confidence in health plan:** Measures patients' confidence in their health plan, their adherence to this plan, and the level of support they have in managing their health-related problems (6 items).

PROM includes evaluations of feasibility, structural validity, internal consistency, construct validity, and responsiveness.^{13,34} In this study, feasibility was assessed by the amount and pattern of missing data, readiness of patients to complete the measure, and by response rate between baseline and follow-up. Structural validity was tested using exploratory factor analysis with principal axis factoring on fully completed questionnaires.³⁵ This is a statistical technique used to reduce a larger number of items into a smaller number of common factors that reflect shared variance.³⁵ The number of factors extracted was decided by a combination of Kaiser's rule (eigenvalues >1), the scree plot,³⁵ and by interpretability of domains. Internal consistency reliability of each domain was assessed using Cronbach's α .²⁵ Construct validity was explored by testing pre-specified hypotheses about the relationships between PCOQ domain scores and comparator PROMs, using Spearman correlations.²⁵ Finally, responsiveness was tested by comparing Glass's Δ ²⁵ for patients expected to improve versus those expected to remain unchanged, based on patient responses to the item asking about change ('Thinking about the main

problem you consulted your GP with at your recent appointment, is this problem: completely better, much better, better, slightly better, same, slightly worse, worse, n/a?'). Glass's Δ is a repeated measures effect size, calculated as the mean change in scores (baseline to follow-up) divided by the standard deviation of scores at baseline.

RESULTS

Feasibility

The PCOQ was accepted by 718 people in the waiting room, and finished by 602 (84%). Missing data at baseline varied between 1% and 7% per item, with 2.5% missing overall. In all, 512 questionnaires (85%) had no missing data. Of the 602 patients completing the PCOQ at baseline, 264 (44%) completed the follow-up questionnaire.

Table 1 shows patient characteristics at baseline and at follow-up. Older patients (apart from the ≥75-year-olds group) had higher response rates at follow-up.

Structural validity

Kaiser's rule suggested a four- or five-factor solution, and the scree plot a two-factor solution. (Further details are available from the author.) Exploratory factor analyses were thus conducted using the complete baseline data ($n = 512$) for solutions ranging from one to five factors. The four-factor solution with oblique (promax) rotation³⁵ provided the most interpretable simple structure (Table 2). Three items that did not load highly on any factor were removed. Two of these items were related to health concerns, a construct also reflected in other items. The third was on medication side effects. The obliquely rotated factors were moderately correlated (0.29–0.51). The factors were labelled and described (Box 1).

A score was calculated for each domain using a simple average of item scores for each domain (scored 1–5). Alternative scores incorporating factor weights were also produced, but converged with the non-weighted scores ($r = 0.99$) and the simpler method was thus preferred. The PCOQ score distributions are shown in Table 3. There is evidence of a ceiling effect in some domains. For example, all patients scoring in the top quintile for both health knowledge and self-care, and confidence in health provision scored at the ceiling. However, the ceiling effect for confidence in health provision was lower than the ceiling for the main comparator, last appointment score (25% as opposed to 36%). Similarly, the ceiling effect for health and wellbeing was lower than the EQ-5D (12% at the ceiling as opposed to 20% for the EQ-5D).

Table 2. Final four-factor solution after promax rotation on baseline data (n = 512)

PCOQ Item loadings ^{a,b}	Health and wellbeing	Health knowledge and self-care	Confidence in health provision	Confidence in health plan	Uniqueness
Q1 Pain	0.58				0.71
Q2 Other physical symptoms	0.61				0.64
Q3 Depression	0.72				0.37
Q4 Stress	0.66				0.42
Q5 Enjoying life	0.85				0.32
Q6 Able to do normal activities	0.84				0.37
Q7 Health concerns	0.72				0.37
Q8 Concerns about serious illness	0.47				0.66
Q9 Confidence clinicians will listen			0.90		0.20
Q10 Confidence clinicians will help			0.95		0.15
Q11 Confidence in clinicians' expertise			0.89		0.25
Q12 Confident clinicians would spot serious illness			0.83		0.29
Q13 Can trust clinicians			0.88		0.19
Q14 Confidence in obtaining good health care when needed			0.78		0.31
Q15 Can prevent health problems		0.77			0.40
Q16 Can stay healthy		0.74			0.45
Q17 Understand illness		0.81			0.34
Q18 Can manage symptoms		0.77			0.32
Q19 Have support to manage in life				0.47	0.54
Q20 Have support to deal with worry				0.49	0.52
Q21 Dealing with cause of illness				0.80	0.30
Q22 On the right path				0.76	0.34
Q23 Adherence to medication				0.48	0.77
Q24 Adherence to lifestyle				0.58	0.72

^aLoadings <0.3 have not been shown. ^bSee Primary Care Outcomes Questionnaire for full questions: <http://www.bristol.ac.uk/primaryhealthcare/resources/pcog/>.

Internal consistency

Cronbach's α was above the standard of 0.70 for each domain³⁴ as follows: health and wellbeing ($\alpha = 0.88$), confidence in health provision ($\alpha = 0.95$), health knowledge and self-care ($\alpha = 0.85$), and confidence in health plan ($\alpha = 0.77$).

Construct validity

Spearman rho correlations of PCOQ domain scores with comparator PROMs (Table 4) indicate convergent and discriminant construct validity:²⁵ domain scores correlate with PROMs as expected on conceptual grounds (for example, health and wellbeing with EQ-5D, $P = 0.75$), and less strongly with the other PROMs that reflect different constructs (for example, health and wellbeing with PAM-13, $P = 0.28$).

Responsiveness

The authors hypothesised that patients responding positively to the change in main problem item ('Slightly better' to 'Much better') would have a positive Glass's Δ

(>0.2) for health and wellbeing, health knowledge and self-care, and confidence in health plan, if they indicated problems in these domains when they attended their GP (patients were excluded if they were at the ceiling at baseline, as this indicated that they had no problem in that domain when attending). The authors also expected patients who were perfectly satisfied on the last appointment score to have a positive Glass's Δ for the domain confidence in health provision. They only used the top score because the last appointment score had a strong positive skew, which meant that using, for example, the top two categories would have included most patients.

The effect sizes for patients expected to improve are positive, and small to moderate in magnitude, with confidence intervals excluding zero (indicating statistical significance at the 0.05 level). Effect sizes for patients expected to stay the same all approach zero (Table 5). The sample of patients reporting a deterioration on

Table 3. PCOQ factor scores (complete baseline data)

	Health and wellbeing		Health knowledge and self-care		Confidence in health provision		Confidence in health plan	
	Mean	95% CI	Mean	95% CI	Mean	95% CI	Mean	95% CI
<i>n</i>	593		594		589		591	
Overall score	3.52	3.45 to 3.60	4.14	4.07 to 4.21	4.02	3.95 to 4.09	4.14	4.08 to 4.19
Standard deviation	0.94		0.83		0.80		0.69	
Breakdown by quintile								
Top quintile	4.76	4.73 to 4.79	5.00	5.00 to 5.00	5.00	5.00 to 5.00	4.97	4.96 to 4.98
Second quintile	4.16	4.13 to 4.19	4.79	4.76 to 4.83	4.59	4.54 to 4.64	4.60	4.59 to 4.62
Third quintile	3.58	3.55 to 3.61	4.26	4.23 to 4.30	4.02	4.01 to 4.03	4.21	4.19 to 4.23
Fourth quintile	2.97	2.93 to 3.00	3.80	3.77 to 3.84	3.70	3.66 to 3.73	3.81	3.79 to 3.84
Bottom quintile	2.14	2.06 to 2.21	2.83	2.72 to 2.94	2.79	2.72 to 2.87	3.08	3.01 to 3.16

the change item was very small and was therefore not analysed.

DISCUSSION

Summary

Following a rigorous process of development, the authors have tested the psychometric properties of a PROM designed to capture outcomes that patients want to obtain from primary care and which doctors seek to deliver. This meets a need for an instrument that can determine the

effects of alternative forms of primary care, where patients have various problems and reasons for consultation.

Strengths and limitations

The PCOQ has advantages over existing PROMs, and meets recommended standards for psychometric testing in this sample of primary care patients.³⁴ Strengths of the study include successful data collection, a simple factor structure with good construct validity, and a prospective

Table 4. Actual and hypothesised correlations of factors with comparator PROMs

	Health and wellbeing	Health knowledge and self-care	Confidence in health provision	Confidence in health plan
Hypothesised correlations	EQ-5D	+ve (high) ^a		+ve (moderate) ^b
	IPQ	-ve (high/moderate) ^{a,b}	-ve (moderate) ^b	-ve (moderate) ^b
	PAM-13		+ve (moderate) ^b	
	Last appointment score		+ve (moderate) ^b	+ve (moderate) ^b
	Recommend GP item		+ve (high) ^a	
	Support for LTCs			+ve (moderate) ^b
	Change in main problem	Correlations above 0.4 not expected		
Actual correlations	EQ-5D	0.75^d	0.23 ^d	0.10
	IPQ	-0.69^d	-0.33^d	-0.19 ^c
	PAM-13	0.28 ^d	0.44^d	0.30 ^d
	Last appointment score	0.19 ^c	0.44^d	0.62^d
	Recommend GP item	0.12 ^c	0.39 ^d	0.53^d
	Support for LTCs	0.42 ^d	0.29 ^d	0.17 ^d
	Change in main problem	0.32 ^d	0.32 ^d	0.27 ^d

Bold entries show hypothesis. ^aHigh = ≥0.6. ^bModerate = ≥0.4. ^cP<0.05. ^dP<0.001. EQ-5D = European Quality of Life-5 Dimensions. IPQ = Illness Perception Questionnaire. LTC = long-term condition. PAM-13 = Patient Activation Measure (13-item version). PROM = Patient Reported Outcome Measure.

Table 5. Effect sizes (Glass's Δ) between PCOQ baseline and follow-up scores

Subset of observations analysed	Health and wellbeing		Health knowledge and self-care		Confidence in health provision		Confidence in health plan	
	Glass's Δ (95% CI)	n	Glass's Δ (95% CI)	n	Glass's Δ (95% CI)	n	Glass's Δ (95% CI)	n
Patients expected to improve ^a	0.52 (0.25 to 0.78)	118	0.37 (0.06 to 0.68)	82	0.48 (0.05 to 0.90)	45	0.27 (0.00 to 0.54)	105
Patients expected to stay the same ^b	0.13 (-0.17 to 0.42)	90	0.14 (-0.19 to 0.47)	71	-0.02 (-0.25 to 0.21)	141	0.09 (-0.23 to 0.41)	74

^aIdentified by the maximum last appointment score for the confidence in health provision domain, and a score >3 (same) in change in main problem for all other domains.

^bIdentified by a score below maximum last appointment score for the confidence in health provision domain, and a score of 3 (same) in change in main problem for all other domains. PCOQ = Primary Care Outcomes Questionnaire.

design that enabled tests of responsiveness.

The study also had limitations. Some patients did not accept a questionnaire, and 16% were called to an appointment before completing the baseline PCOQ. Though efforts were made to obtain a mix of urban and rural practices, with different deprivation scores, data were collected in a relatively small number of practices. The sample was representative of general practice consultation in terms of sex³⁶ and number of long-term conditions.¹⁶ However, patients >75 years were slightly under-represented,³⁷ perhaps because housebound patients were excluded, or more declined to take part. Response rates from ethnic groups may have been affected by the lack of interpretation facilities. Only 44% of patients returned the follow-up form, and this differed systematically by age (though not by sex, long-term conditions, or ethnicity). This response rate is comparable to similar postal questionnaires³⁸ and was anticipated by the authors' protocol. Baseline and follow-up scores were compared only for the final responsiveness tests. Furthermore, this responsiveness analysis was based on comparing patients who were separated into two groups based on a change score, and there is no reason to assume different numbers of non-responders across groups. However, the authors recognise that selection bias may have been introduced. As with comparable questionnaires,^{39,40} most items and domain scores were positively skewed. The factor solution had high uniqueness for some items. This can indicate that the item is not strongly related to others,³⁵ but, because of the important content of these variables (for example, pain and adherence), the authors chose to include them. The Cronbach's α of 0.95 for the confidence in health provision factor may indicate some item redundancy,²⁵ and future validation might further reduce the items. A final limitation

relates to the lack of a gold standard for measuring change. The change in main problem and last appointment score were used as proxies, but these are imperfect measures. However, this is a necessary limitation of developing a PROM with a new and unique set of constructs, and the evidence for responsiveness is akin to the concept of 'construct responsiveness' as described by the COSMIN (COnsensus-based Standards for the selection of health Measurement INstruments) group.^{41,42}

Comparison with existing literature

Taken together, the four domains of the PCOQ have similarities with the concept of health capability, defined as combining health agency (an individual's ability to achieve health goals and act as agents of their own health) and health functioning (the outcome of actions to maintain or improve health).⁴³ The four domains are scored separately, and each domain has benefits over other generic PROMs.

The health and wellbeing domain, which includes physical and/or emotional symptoms, life effects, and health concerns, was the most responsive, and had a lower ceiling effect than the EQ-5D, another measure of health status. However, the authors did not carry out a head-to-head comparison of responsiveness, and this is a future research requirement. Similarly, the responsiveness of the health knowledge and self-care domain should be compared with similar measures such as the PAM-13.³⁰ Although it showed a stronger ceiling effect than the PAM-13, this domain includes areas often receptive to intervention, such as patient understanding of health problems, while excluding areas, captured by PAM-13, that may be less responsive, such as figuring out solutions to new health problems. The confidence in health plan domain refers to patients trusting and following their health plan,

and having support to enable this. This is a broad construct that subsumes different concepts, and no existing PROM is readily comparable. Lastly, the confidence in health provision domain includes concepts similar to those normally found in Patient-Reported Experience Measures (PREMs) as opposed to PROMs, such as whether the clinician listens.³¹ However, unlike a PREM, which refers to perceptions of a particular consultation, the PCOQ refers to patients' current levels of confidence in these aspects of their health providers. These are best viewed as outcomes rather than experiences. That these levels are amenable to change following a consultation is a valuable result of this study.

Implications for research and practice

This study has demonstrated that the PCOQ is valid, internally consistent, and

responsive among this sample of primary care patients. Because this study represents the first validation of the PCOQ, the authors recommend usage alongside other PROMs until properties are confirmed. The PCOQ was specifically developed to test the benefits of service-level interventions in primary care, and thus fills an important gap in the literature. The alternative to the PCOQ is multiple instruments measuring four different constructs. This would not only require a longer questionnaire, but also usage of instruments not designed to measure outcomes that primary care patients seek. No existing PROM covers the PCOQ unique construct. It therefore offers a timely opportunity to enhance research and policymaking in primary care during a period of high demand for new interventions in this area.

Funding

This study was funded by the National Institute for Health Research School for Primary Care Research (NIHR SPCR), which is a partnership between the universities of Bristol, Cambridge, Keele, Manchester, Newcastle, Nottingham, Oxford, Southampton, and University College London. The views expressed are those of the authors and not necessarily those of the NHS, the NIHR, or the Department of Health. Mairead Murphy was partly funded by the Avon Primary Care Research Collaborative.

Ethical approval

Ethical approval was granted through the National Health Service Research Ethics Committee process. Approval was granted by London Central Committee, reference 14/LO/2036.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

The authors have declared no competing interests.

Acknowledgements

The authors thank all the participants in this study, in particular the patients who completed the questionnaires and the primary care centres that allowed it to be administered. The authors also thank the Bristol Primary Care Research Network for assisting with recruiting the participants, and the NIHR SPCR for funding the research.

Open access

This article is Open Access: CC BY-NC 4.0 licence (<http://creativecommons.org/licenses/by-nc/4.0/>).

Discuss this article

Contribute and read comments about this article: bjgp.org/letters

REFERENCES

1. Salisbury C, Goodall S, Montgomery AA, *et al*. Does Advanced Access improve access to primary health care? Questionnaire survey of patients. *Br J Gen Pract* 2007; **57(541)**: 615–621.
2. O’Cathain A, Coster J, Salisbury C, *et al*. Do walk-in centres for commuters work? A mixed methods evaluation. *Br J Gen Pract* 2009; <https://doi.org/10.3399/bjgp09X473150>.
3. Campbell JL, Fletcher E, Britten N, *et al*. Telephone triage for management of same-day consultation requests in general practice (the ESTEEM trial): a cluster-randomised controlled trial and cost-consequence analysis. *Lancet* 2014; **384(9957)**: 1859–1868.
4. Jackson GL, Williams JW Jr. Does PCMH ‘work’? The need to use implementation science to make sense of conflicting results. *JAMA Intern Med* 2015; **175(8)**: 1369–1370.
5. Nielsen M, Buelt L, Patel K, Nichols LM. *The patient-centered medical home’s impact on cost and quality: annual review of evidence 2014–2015*. Washington, DC: Patient-Centred Primary Care Collaborative, 2016.
6. Friedberg MW, Rosenthal MB, Werner RM, *et al*. Effects of a medical home and shared savings intervention on quality and utilization of care. *JAMA Intern Med* 2015; **175(8)**: 1362–1368.
7. Fortin M, Bravo G, Hudon C, *et al*. Prevalence of multimorbidity among adults seen in family practice. *Ann Fam Med* 2005; **3(3)**: 223–228.
8. Olayiwola JN, Anderson D, Jepeal N, *et al*. Electronic consultations to improve the primary care–specialty care interface for cardiology in the medically underserved: a cluster-randomized controlled trial. *Ann Fam Med* 2016; **14(2)**: 133–140.
9. Sharma AE, Willard-Grace R, Hessler D, *et al*. What happens after health coaching? Observational study 1 year following a randomized controlled trial. *Ann Fam Med* 2016; **14(3)**: 200–207.
10. Hudon C, Chouinard MC, Diadiou F, *et al*. Case management in primary care for frequent users of health care services with chronic diseases: a qualitative study of patient and family experience. *Ann Fam Med* 2015; **13(6)**: 523–528.
11. Black N. Patient reported outcome measures could help transform healthcare. *BMJ* 2013; **346**: f167.
12. Heath I, Rubinstein A, Stange KC, van Driel ML. Quality in primary health care: a multidimensional approach to complexity. *BMJ* 2009; **338**: b1242.
13. Mokkink LB, Terwee CB, Patrick DL, *et al*. The COSMIN study reached international consensus on taxonomy, terminology, and definitions of measurement properties for health-related patient-reported outcomes. *J Clin Epidemiol* 2010; **63(7)**: 737–745.
14. Salisbury C, Procter S, Stewart K, *et al*. The content of general practice consultations: cross-sectional study based on video recordings. *Br J Gen Pract* 2013; <https://doi.org/10.3399/bjgp13X674431>.
15. 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(9836)**: 37–43.
16. Salisbury C, Johnson L, Purdy S, *et al*. Epidemiology and impact of multimorbidity in primary care: a retrospective cohort study. *Br J Gen Pract* 2011; <https://doi.org/10.3399/bjgp11X548929>.
17. Ware JE Jr, Sherbourne CD. The MOS 36-item short-form health survey (SF-36). I. Conceptual framework and item selection. *Med Care* 1992; **30(6)**: 473–483.
18. Brooks R. EuroQol: the current state of play. *Health Policy* 1996; **37(1)**: 53–72.
19. Venning P, Durie A, Roland M, *et al*. Randomised controlled trial comparing cost effectiveness of general practitioners and nurse practitioners in primary care. *BMJ* 2000; **320(7241)**: 1048–1053.
20. McKinley RK, Cragg DK, Hastings AM, *et al*. Comparison of out-of-hours care provided by patients’ own general practitioners and commercial deputising services: a randomised controlled trial. II: The outcome of care. *BMJ* 1997; **314(7075)**: 190–193.
21. Paterson C. Measuring outcomes in primary care: a patient generated measure, MYMOP, compared with the SF-36 health survey. *BMJ* 1996; **312(7037)**: 1016–1020.
22. Howie JG, Heaney DJ, Maxwell M, *et al*. Quality at general practice consultations: cross-sectional survey. *BMJ* 1999; **319(7212)**: 738–743.
23. Haddad S, Potvin L, Roberge D, *et al*. Patient perception of quality following a visit to a doctor in a primary care unit. *Fam Pract* 2000; **17(1)**: 21–29.
24. Valderas JM, Fitzpatrick R, Roland M. Using health status to measure NHS performance: another step into the dark for the health reform in England. *BMJ Qual Saf* 2012; **21(4)**: 352–353.
25. Streiner DL, Norman GR. *Health measurement scales: a practical guide to their development and use*. 4th edn. New York, NY: Oxford University Press, 2008.
26. Murphy M, Hollinghurst S, Turner K, Salisbury C. Patient and practitioners’ views on the most important outcomes arising from primary care consultations: a qualitative study. *BMC Fam Pract* 2015; **16(108)**. DOI: 10.1186/s12875-015-0323-9.
27. Murphy M, Hollinghurst S, Salisbury C. Agreeing the content of a patient-reported outcome measure for primary care: a Delphi consensus study. *Health Expect* 2017; **20(2)**: 335–348. DOI: 10.1111/hex.12462.
28. Murphy M, Hollinghurst S, Salisbury C. Qualitative assessment of the primary care outcomes questionnaire: a cognitive interview study. *BMC Health Serv Res* 2018; DOI: 10.1186/s12913-018-2867-6.
29. Herdman M, Gudex C, Lloyd A, *et al*. Development and preliminary testing of the new five-level version of EQ-5D (EQ-5D-5L). *Qual Life Res* 2011; **20(10)**: 1727–1736.
30. Hibbard JH, Mahoney ER, Stockard J, Tusler M. Development and testing of a short form of the patient activation measure. *Health Serv Res* 2005; **40(6 Pt 1)**: 1918–1930.
31. 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.
32. Broadbent E, Petrie KJ, Main J, Weinman J. The brief illness perception questionnaire. *J Psychosom Res* 2006; **60(6)**: 631–637.
33. Kamper SJ, Maher CG, Mackay G. Global rating of change scales: a review of strengths and weaknesses and considerations for design. *J Man Manip Ther* 2009; **17(3)**: 163–170.
34. Reeve BB, Wyrwich KW, Wu AW, *et al*. ISOQOL recommends minimum standards for patient-reported outcome measures used in patient-centered outcomes and comparative effectiveness research. *Qual Life Res* 2013; **22(8)**: 1889–1905.
35. Tabachnick BG, Fidell LS, eds. *Using multivariate statistics*. Harlow: Pearson, 2013.
36. Wang Y, Hunt K, Nazareth I, *et al*. Do men consult less than women? An analysis of routinely collected UK general practice data. *BMJ Open* 2013; **3(8)**: e003320.
37. Hobbs FD, Bankhead C, Mukhtar T, *et al*. Clinical workload in UK primary care: a retrospective analysis of 100 million consultations in England, 2007–14. *Lancet* 2016; **387(10035)**: 2323–2330.
38. Croker JE, Swancutt DR, Roberts MJ, *et al*. Factors affecting patients’ trust and confidence in GPs: evidence from the English national GP patient survey. *BMJ Open* 2013; **3(5)**: e002762.
39. McDowell I. *Measuring health: a guide to rating scales and questionnaires*. 3rd edn. New York, NY: Oxford University Press, 2006.
40. Bowling A. *Measuring health: a review of quality of life measurement scales*. 3rd edn. Maidenhead: Open University Press, 2004.
41. Mokkink LB, Terwee CB, Knol DL, *et al*. The COSMIN checklist for evaluating the methodological quality of studies on measurement properties: a clarification of its content. *BMC Med Res Methodol* 2010; **10**: 22.
42. Angst F. The new COSMIN guidelines confront traditional concepts of responsiveness (and author response). *BMC Med Res Methodol* 2011; **11**: 152.
43. Ruger JP. Health capability: conceptualization and operationalization. *Am J Pub Health* 2010; **100(1)**: 41–49.