ORIGINAL ARTICLE

Capturing users' experience of UK out-of-hours primary medical care: piloting and psychometric properties of the Out-of-hours Patient Questionnaire

J L Campbell, A Dickens, S H Richards, P Pound, M Greco, P Bower

Supplementary material available at http:// qshc.bmj.com/supplemental

See end of article for authors' affiliations

Correspondence to:
J L Campbell, General Practice and Primary Care, Peninsula Medical School, St Luke's Campus, Smeall Building, Magdalen Road, Exeter EX1 2LU, UK; john.campbell@pms.ac.uk

Accepted 27 May 2007

Background and objective: Provision of out-of-hours care in the UK National Health Service (NHS) has changed in recent years with new models of provision and the introduction of national quality requirements. Existing survey instruments tend to focus on users' satisfaction with service provision; most were developed without undertaking supporting qualitative fieldwork. In this study, a survey instrument was developed taking account of these changes in service provision and undertaking supporting qualitative fieldwork. This paper reports on the development and psychometric properties of the new survey instrument, the Out-of-hours Patient Questionnaire (OPQ), which aims to capture information on the entirety of users' experiences of out-of-hours care, from the decision to make contact through to completion of their care management.

Methods: An iterative approach was undertaken to develop the new instrument which was then tested in users of out-of-hours services in three geographically distributed UK settings. For the purposes of this study, "service users" were defined as "individuals about whom contact was made with an out-of-hours primary care medical service", whether that contact was made by the user themselves, or via a third party. Analysis was undertaken of the acceptability, reliability and validity of the survey instrument.

Results: The OPQ tested is a 56-item questionnaire, which was distributed to 1250 service users. Respondents were similar in respect of gender, but were older and more affluent (using a proxy measure) than non-respondents. Item completion rates were acceptable. Respondents sometimes completed sections of the questionnaire which did not equate to their principal mode of management as recorded in the record of the contact. Preliminary evidence suggests the OPQ is a valid and reliable instrument which contains within it two discrete scales—a consultation satisfaction scale (nine items) and an "entry-access" scale (four items). Further work is required to determine the generalisability of findings obtained following use of the OPQ, especially to non-white user populations.

Conclusion: The OPQ is an acceptable instrument for capturing information on users' experiences of out-of-hours care. Preliminary evidence suggests it is both valid and reliable in use. Further work will report on its utility in informing out-of-hours service planning and configuration and standard-setting in relation to UK national quality requirements.

Amongst Western industrialised nations, heterogeneity exists in both the social and legal obligations and in the nature and requirements of national health systems towards requirements for 24-h coverage of care.[1] Where primary care practitioners have responsibility of out-of-hours care provision, broadly, four models of involvement exist[2]:

- personal care by family doctors (increasingly uncommon);
- out-of-hours rota of family doctors (eg, Germany);
- deputising services (eg, Australia, Spain);
- out-of-hours cooperatives (eg, Denmark).

Some countries fund separate systems of out-of-hours care (eg, Italy), and for many patients (eg, USA) out-of-hours care is provided through hospital-based facilities.

Recent years have seen new models of out-of-hours primary care established in the UK. "Out-of-hours" care is defined as care provided between 18:30 hours and 08:00 hours or at the weekends and public holidays. New approaches include multiple points of access to care, loss of 24-h responsibility for general practitioners for the care of their patients and the establishment of a range of organisations providing out-of-hours services to the National Health Service (NHS) on a contractual basis. The NHS has identified that user experience

and patient need should be at the heart of planning NHS service provision. The introduction of *National quality requirements in the delivery of out-of-hours services*[3] requires providers to regularly audit a sample of patients' experiences of the service.

The UK has experimented with all of these models of care, and currently has a mixed economy of provision, with several recent keynote reviews of current models.[3–5] Primary care trusts are health management organisations coordinating care provision for local populations and receiving about 75% of the NHS budget (http://www.nhs.uk). Since 2004, these trusts have had the responsibility for commissioning out-of-hours care using a variety of provider organisations and health professionals including nurse and emergency care practitioner services as well as family doctor provision, telephone helplines, treatment and walk-in centres, including hospital-based emergency centres, minor injury units and pharmacy-based services.

Although several survey tools have already been developed to assess patient satisfaction with out-of-hours primary medical care services in the UK setting, the only indepth, qualitative study to obtain user views took place in the 1990s,[6] pre-dating major policy and practice changes in this area. Although other

Abbreviations: NHS, National Health Service; OPQ, Out-of-hours Patient Questionnaire; PCA, principal components analysis

patient satisfaction tools developed subsequently[7–9] have used extensive piloting, literature review work, and psychometric testing to evaluate their performance, none has replicated the earlier indepth qualitative work with service users. Given major UK reforms in the delivery of out-of-hours primary care, there is an urgent need to develop and refine existing methods of ascertaining patients' experiences of, and satisfaction with, services by actively involving today's service users.[10] The research described in this paper was undertaken in response to a request from the UK Department of Health to address this need.

In this study, we involved service users in the development of an evidence-based survey instrument for the assessment of patients' experience of out-of-hours primary care. Having previously reported on the qualitative fieldwork undertaken to identify patients' views and priorities for the delivery of out-of-hours care,[11] here we report on the development and psychometric properties of the Out-of-hours Patient Questionnaire (OPQ). The utility of the questionnaire in examining systematic variation between service providers and in informing standard-setting will be reported elsewhere.

METHOD
Development of questionnaire
We developed the OPQ by adopting standard approaches to questionnaire development,[12] drawing on reviews of relevant literature, consultation with experts and piloting of drafts of the questionnaire. The presentation and wording of questionnaire items drew on the expertise of research team members who had contributed to the development of two validated questionnaires recognised and adopted by the UK Department of Health for research and service quality monitoring purposes.[13–15] We initially developed three audit questionnaires, one for use in each of the three main management categories encountered in out-of-hours care (home visit, treatment centre attendance, telephone management), and drawing predominantly on the literature and currently available instruments.[6 7 16] Questionnaire items seeking service user views around key areas of the national quality requirements relating to out-of-hours care were also included.

With a view to extending the generalisability of results, the three sampling sites selected for each element of the fieldwork (audit, qualitative and quantitative) covered a combined population of approximately 2.5 million people, with diverse community profiles in respect of urban/rural mix, ethnicity and demography. The model of out-of-hours care consisted of service users making initial telephone contact with a provider, followed by a preliminary assessment and possible final management of contacts of a purely administrative nature (eg, enquiries regarding surgery opening hours) by a trained call handler. For all other instances, the service user was assessed on the telephone by a clinician regarding the nature/urgency of the contact prior to receiving one of three potential management options:

- completion of telephone management;
- service user invited to attend a local treatment centre;
- service user offered a home visit.

A preliminary audit in three locations (Devon, Kent, Sheffield) identified rather low response rates, which seemed in part to be due to difficulties encountered in sampling by ''management'' option. Some patients with complex health problems may use a service more than once within the sampling frame and receive several different management outcomes within an episode of care, raising the issue as to which type of questionnaire to send. Following our experience in this audit, we therefore elected to develop one instrument

suitable for use in capturing service user views which was flexible enough to capture the entirety of their experience.

While the preliminary audit was underway, the qualitative field work (reported elsewhere[17]) was undertaken using focus group methodology with service users from three locations (Cornwall, Devon, Sheffield) to identify their experiences and priorities for assessing the quality of out-of-hours care. These findings, combined with priority areas identified in the national quality requirements, were used to draw up a comprehensive list of items covering all aspects of the patient experience, organised in a manner consistent with the patient's care pathway. Items from the audit questionnaires which mapped on to patient priorities were selected for inclusion in the OPQ. These 61 items were then piloted with 150 service users from one provider site, with resulting minor changes made to the wording of ambiguous items and five items were dropped as they overlapped substantially with other items (whose response rates were substantially better). This version was reviewed by a panel comprising experts from out-of-hours services and by six user focus group members who had expressed particular interest in being involved in the future development of the questionnaire. All commented on the questionnaire to check for relevance and clarity; following minor changes in wording and presentation, we developed the version of the OPQ evaluated in this paper. The inclusion of focus group members in this process assessed and enhanced the face and content validity of the questionnaire.

Description of OPQ and sampling
The OPQ investigated here comprises 56 items (table 1) presented in eight sections (see appendix and web table 1, available at http://qshc.bmj.com/supplemental). No formal sample size calculation was conducted, but it was decided to collect data from about 500 individuals to allow reliable estimation from multivariate techniques such as factor analysis. Drawing on the experience of local out-of-hours providers, and adopting a conservative estimate of response of 40%, we proposed to survey 1250 individuals across the three geographical areas. We deliberately over-sampled to allow analysis of questionnaire performance for subgroups of individuals receiving the three principal management options. A flow diagram of the access arrangements pertaining in the three sampling sites is included (see web fig 1, available at http://qshc.bmj.com/supplemental).

Sampling took place following out-of-hours contacts in September 2005. Contact details were extracted from the provider dataset for a consecutive sample of 750 users at each

Table 1 Structure of the Out-of-hours Patient Questionnaire (56 items)

		Stem items (n)		Response-dependent items* (n)	
	Section	Report	Evaluative	Report	Evaluative
A	Making contact	5	2	1	2
B	Outcome of call	1	0	1	0
C	Consultation	7	10	3	1
D†	Home visit	3	0	0	1
E†	Treatment centre visit	4	1	0	2
F†	Telephone advice	1	1	0	0
G‡	Overall satisfaction outcomes	0	3	0	0
H	Demographics	7	0	0	0
	Total	28	17	5	6

*Item completion dependent on response to preceding stem item.
†Management specific.
‡Global assessment.

of three sites and a variety of exclusion criteria were applied (see web table 2, available at http://qshc.bmj.com/supplemental). After all exclusions were applied, 1578 patients were potentially eligible. A consecutive sample of the 1250 most recent calls were then included and a numbered questionnaire, information sheet, and reply-paid envelope were sent to the patient (or parent or guardian if the patient was a child) by the service provider within two weeks of the contact. A reminder questionnaire was sent to non-responders after two weeks. Test-retest reliability was assessed in a sample of the first 99 users to return the questionnaire following the first mailing. No reminder was sent to non-respondents in the retest process.

Analysis
Acceptability
The acceptability of the OPQ was evaluated through an examination of overall questionnaire and item response rates, and the distribution of responses across response categories. The age, gender, socioeconomic status (using the individuals' postcode to derive their Townsend index[18] as a proxy measure) and the principal management of the contact (from medical record) of respondents and non-respondents were compared.

Validity
To examine the construct validity of the OPQ, an exploratory principal components analysis was conducted of the 15 attitudinal items which were not management specific or global assessments (table 2).

Inspection of eigenvalues, a scree plot, and item loading patterns was used to explore the underlying structure of responses. In line with best practice,[19] principal conclusions were based on an interpretation of the unrotated factor loadings. The resulting components were interpreted in the light of contributing item content, and used to identify any scales which might exist within the attitudinal components of the questionnaire. Scale scores were calculated for principal components identified (see web table 3, available at http://qshc.bmj.com/supplemental).

The scaling assumptions of the OPQ were tested by examining the correlations between each item and its own scale and its correlation with the other scales, and by examining the distribution of scale scores. Each item should be more correlated to its own scale (item convergent validity) than to the other scales (item discriminant validity).[20] We examined differences in derived scores for two scales identified during the principal components analysis (using t tests) between respondents who reported "excellent" satisfaction with the care provided, and those who reported less than "excellent" satisfaction. In line with recognised predictors of user experience of care,[7 21–23] scale scores were also compared for groups of users defined by their age (four categories) or their reported consultation length (three categories) using analysis of variance.

Reliability
Internal consistency was measured using Cronbach α and item-total correlations. Test-retest reliability was analysed using intra-class correlations between derived scale scores from questionnaires completed by the same respondents two weeks later.

All data manipulation and analysis was conducted using SPSS version 13.0.

RESULTS
Acceptability of OPQ
Response rates
Completed questionnaires were returned by 570/1250 (45.6%) respondents. Respondents were older and more affluent than non-respondents (mean (SD) age = 45.2 (27.3) years versus 36.0 (26.6) years, mean difference (95% CI) = 9.2 (6.1 to 12.2);

mean (SD) Townsend score 0.50 (2.55) versus 0.90 (2.69), mean difference (95% CI) = −0.4 (−0.646 to −0.154)). However, the respondents did not differ significantly from the non-respondents in respect of gender (60.2 versus 59.2% female, respectively; χ^2 0.13, df 1, p = 0.72) and management of the contact (data not presented). Ethnic status of "white" was reported by 452/464 (97.4%) of survey respondents and ranged from 95.9% to 98.7% in the three areas surveyed, compared with a reported white population of 91.2–99.3% in those areas (National Statistics Online, 2006 http://neighbourhood.statistics.gov.uk/dissemination).

Respondents were predominantly of "white" ethnicity and reported higher levels of limiting longstanding illnesses compared with UK national statistics[24] (97.4% versus 90.9%, 45.0% versus 33.6%, respectively), whereas the proportion of owner occupiers was similar (65.1% versus 68.7%, respectively) and people in full or part-time employment were somewhat under-represented (30.9% versus 52.6%, respectively).

Only 3/570 (0.5%) subjects completed fewer than 50% of the 38 items which were not management specific or which provided reasons for the directly preceding item. Four of these 38 items (which health professional conducted the consultation; ethnic group; occupation; were you happy with [the final management of your call]) had missing values exceeding 10% of responses (maximum 20.4%). Table 2 presents the distribution of responses to attitudinal items.

Section completion
Between 78.8% and 95.2% of respondents appropriately completed sections of the survey related to specific management (eg, relating to attendance at a treatment centre). However, some confusion in responses was evident from inappropriate completion of sections of the questionnaire. For example, of individuals who reported receiving a home visit, about 35% completed questionnaire sections relating to telephone advice and 12% completed questionnaire sections relating to treatment centre attendance. Also, around 29% of users who reported attending a treatment centre completed questionnaire sections relating to receiving telephone advice.

There was also some discrepancy evident in patient reports of the last management option provided by the service when compared with the medical record. Taking the patient report as the denominator, the crude agreement for home visiting, for treatment centre appointment and for telephone advice only was 89/120 (74.2%), 239/291 (82.1%) and 124/132 (93.9%), respectively.

Validity
Preliminary evidence of content validity of the OPQ was derived from indepth empirical work with service users, and subsequent feedback from service users and service managers supporting questionnaire development.[11] Further evidence derives from the principal components analysis (PCA).

Principal components analysis
PCA revealed the presence of two components with eigenvalues exceeding 1, accounting for 68.1% of the variance (54.7%, and 13.4%, respectively). An inspection of the scree plot (fig 1) revealed a clear break after the second component, and on this basis it was decided to retain two components for further investigation. While inspection of factor loadings (table 3) identified some overlap in loading patterns in respect of the first four items, these items generally loaded most substantially on the second component, and intuitively seemed to relate together in respect of item content. One item (availability of the medical history), with only modest loadings on either factor, was excluded.

Table 2 Out-of-hours Patient Questionnaire: item completion of 15 attitudinal items which were not management-specific or global assessments

	Very poor n (%)	Poor n (%)	Acceptable n (%)	Good n (%)	Excellent n (%)	N/A n (%)	Missing n (%)	Total N
How do you rate the time it took for your call to be answered?	4 (0.7)	11 (1.9)	112 (19.6)	207 (36.3)	229 (40.2)	–	7 (1.2)	570
Please rate the helpfulness of the call operator	2 (0.4)	9 (1.6)	68 (11.9)	246 (43.2)	235 (41.2)	–	10 (1.8)	570
Please rate the extent to which you felt the call operator listened to you	2 (0.4)	14 (2.5)	64 (11.2)	215 (37.7)	221 (38.8)	–	54 (9.5)	570
How do you rate the time it took for a health professional to call you back?	7 (1.5)	32 (6.8)	108 (22.9)	137 (29.0)	157 (33.3)	–	31 (6.6)	472*
How do you rate the length of your consultation with the health professional?	14 (2.5)	23 (4.0)	116 (20.4)	223 (39.1)	170 (29.8)	–	24 (4.2)	570
Rate the thoroughness of the consultation	18 (3.2)	22 (3.9)	75 (13.2)	228 (40.0)	201 (35.3)	6 (1.1)	20 (3.5)	570
Rate the accuracy of the diagnosis	15 (2.6)	24 (4.2)	72 (12.6)	211 (37.0)	189 (33.2)	30 (5.3)	29 (5.1)	570
Rate the treatment you were given	23 (4.0)	19 (3.3)	61 (10.7)	175 (30.7)	168 (29.5)	87 (15.3)	37 (6.5)	570
Rate the advice and information you were given	20 (3.5)	33 (5.8)	63 (11.1)	229 (40.2)	196 (34.4)	10 (1.8)	19 (3.3)	570
Rate the warmth of the health professional's manner	20 (3.5)	15 (2.6)	75 (13.2)	186 (32.6)	251 (44.0)	6 (1.1)	17 (3.0)	570
Rate the extent to which you felt listened to	20 (3.5)	27 (4.7)	65 (11.4)	204 (35.8)	236 (41.4)	4 (0.7)	14 (2.5)	570
Rate the extent to which you felt things were explained to you	20 (3.5)	28 (4.9)	65 (11.4)	208 (36.5)	217 (38.1)	13 (2.3)	19 (3.3)	570
Rate the respect you were shown	19 (3.3)	15 (2.6)	59 (10.4)	190 (33.3)	268 (47.0)	5 (0.9)	14 (2.5)	570

	Definitely not n (%)	Possibly not n (%)	Not sure n (%)	Yes, possibly n (%)	Yes, definitely n (%)		Missing n (%)	Total N
Do you think the out-of-hours service knew enough about your medical history?	61 (10.7)	129 (22.6)	204 (35.8)	104 (18.2)	60 (10.5)		12 (2.1)	570

	Yes, I was happy n (%)	No, I should have had a home visit n (%)	No, I should have been seen at a treatment centre n (%)	No, I should have been given advice on the telephone n (%)	Other (specify) n (%)	Missing n (%)	Total N
Were you happy with [the management offered]?	408 (71.6)	33 (5.8)	6 (1.1)	7 (1.2)	0 (0.0)	116 (20.4)	570 (100.0)

*Denominator varies since not all patients received a health professional call back.
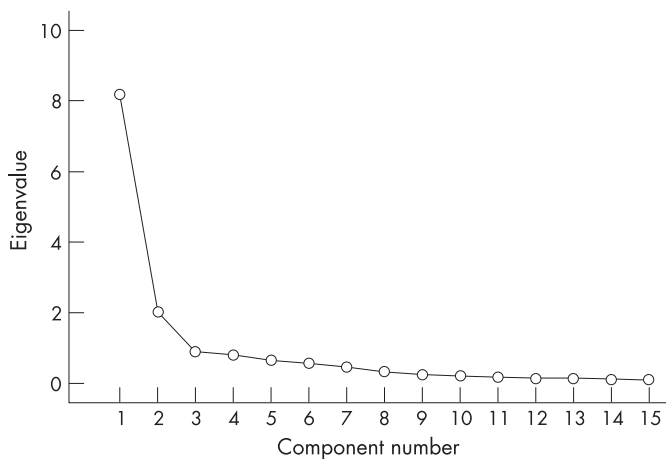
Figure 1 Scree plot of eigenvalues for 15 attitudinal variables in the Out-of-hours Patient Questionnaire.

The interpretation of the two principal components identified suggested that the first related predominantly to issues concerning the interaction between user and the health professional. Of the 10 items loading on this component, nine (items 5–13, table 3) fitted this interpretation; the tenth appeared somewhat different, being a dichotomised variable relating to users' satisfaction with the management option provided (whether a home visit, treatment centre attendance or telephone advice). This item was also highly correlated[25] (r = −0.54, p<0.001) with the global item relating to satisfaction with care, and was therefore omitted. The remaining nine items contributed to the "consultation satisfaction scale". The second component encompassed four variables (items 1–4, table 3), all of which related to users' initial experience of contact with the out-of-hours service (addressing such issues as telephone accessibility, and the performance of reception staff at the time of the initial contact). This scale was labelled the "entry access scale". The interpretations provided here are consistent with a recently proposed model of the accessibility of primary care services in which distinction has been drawn between the separate elements of "entry" access and "in-system" access.[26]

Additional tests of construct validity

Across the sample, scores for the two scales were moderately correlated[25] (r = 0.43, p<0.001) providing evidence of convergence in the domains addressed. Respondents who reported "excellent" levels of global satisfaction with their care scored significantly higher on each of the scales than those who reported lower levels of satisfaction (consultation satisfaction mean (SD) scale score = 90.8 (12.5) versus 68.2 (22.2), mean difference (95% CI) = −22.6 (−25.5 to −19.6); entry access mean scale score = 89.8 (12.3) versus 72.4 (16.9), mean difference (95% CI) = −17.4 (−19.9 to −14.9)). Older users and users reporting longer consultations had higher mean consultation satisfaction scores than younger users, and users reporting receiving shorter consultations (table 4). Entry access scale scores were similar for both groups of respondents.

Reliability

Both the consultation satisfaction scale and the entry access scale had evidence of high internal consistency with α (table 5) being in excess of the 0.70 recommended for group comparison.[27] Both had acceptable homogeneity, with inter-item and item-total correlations exceeding 0.2.[12] The α, if individual contributing items were deleted, was constant at 0.96 across all items for the consultation scale. In both scales there was some evidence of a floor effect with only small numbers scoring the lowest possible, or lowest derived scores.

Subjects completing retest questionnaires (n = 28) were similar (p>0.05) to non-responders (n = 71) in respect of age (mean = 52.1 (22.8) years versus 43.5 (27.1), mean difference (95% CI) = −8.6 (−20.1 to 2.9)), gender (60.7% versus 54.9%

Table 4 Consultation satisfaction and entry access scale scores for groups of users defined by age and the reported consultation length

	Consultation satisfaction (mean scale score, SE)	F	Entry access (mean scale score, SE)	F
Age (years, n)				
0–4 (57)	70.0 (3.3)	5.6*	76.9 (2.9)	0.8
5–11 (43)	80.0 (2.5)		82.3 (2.2)	
16–64 (299)	74.8 (1.4)		78.5 (1.1)	
65+ (158)	81.7 (1.3)		79.1 (1.2)	
Reported consultation length (length, n)				
<10 minutes (282)	69.0 (1.5)	40.9**	77.6 (1.1)	1.5
10–20 minutes (199)	83.5 (1.1)		80.4 (1.2)	
>20 minutes (70)	88.2 (1.8)		79.2 (2.1)	

*p<0.01
**p<0.001.

Table 3 Out-of-hours Patient Questionnaire: matrix of unrotated item factor loadings

	Component	
	1	2
How do you rate the time it took for your call to be answered?	0.412	0.663
Please rate the helpfulness of the call operator	0.495	0.732
Please rate the extent to which you felt the call operator listened to you	0.505	0.703
How do you rate the time it took for a health professional to call you back?	0.531	0.502
Dichotomised rating satisfaction with disposal	−0.607	
How do you rate the length of your consultation with the health professional?	0.830	
Rate the thoroughness of the consultation	0.879	
Rate the accuracy of the diagnosis	0.836	
Rate the treatment you were given	0.858	
Rate the advice and information you were given	0.884	
Rate the warmth of the health professional's manner	0.855	
Rate the extent to which you felt listened to	0.901	
Rate the extent to which you felt things were explained to you	0.889	
Rate the respect you were shown	0.845	
Do you think the out-of-hours service knew enough about your medical history?	0.401	

Table 5 Reliability of the Out-of-hours Patient Questionnaire

	α	Inter-item correlation (range)	Item-total correlations (range)	Item deletion (α, range)	Scale score (mean %, (95% CI))	Lowest score possible* (% of sample)	Highest score possible* (% of sample)
Consultation satisfaction (nine items)	0.96	0.63–0.89	0.77–0.90	0.96–0.96	76.7(74.7 to 78.4)	0.5	15.6
Entry access (four items)	0.82	0.45–0.86	0.56–0.73	0.73–0.82	78.9 (77.4 to 80.3)	0.0	23.3

*Among individuals with at least four valid responses across all nine items within the consultation scale (n = 400), or two valid responses out of the four potential items contributing to the entry access scale (n = 563).

female, χ^2 = 0.27, p = 0.60)) and Townsend score (mean = −0.6 (2.1) versus 0.1 (2.8), mean difference (95% CI) = −0.7 (−1.67 to 0.27)). The test-retest correlation of consultation satisfaction scale scores was 0.76 (p<0.001), whereas the retest correlation for the entry access scale score was 0.60 (p<0.01), only slightly less than the recommended 0.70 criteria (Scientific Advisory Committee of the Medical Outcomes Trust 1995).

A final version of the questionnaire is available on the web (http://www.pms.ac.uk/pms/research/outOfHours.php).

DISCUSSION
Context
This study focused on service users' experience of out-of-hours care provided by three care organisations. Users almost invariably contacted the service by telephone, having some initial interaction with (non-clinical) call handlers, before being called back by a healthcare professional who undertook clinical triage before offering a management option. Users may have had multiple contacts with the out-of-hours service over a short period, each with different patterns of service use. Capturing the complexity of this interaction within a questionnaire was a challenging undertaking.

Sampling was carried out across sites with a diverse range of social demographic and geographic characteristics, which we believe are broadly typical of the UK population. Analysis of pooled data was undertaken and this may be associated with some loss of the inevitable heterogeneity of individual user views.

Acceptability
The OPQ was acceptable to service users as evidenced by satisfactory survey response rates and item-completion rates, and by feedback from users on the clarity of questions, on legibility, and on other presentational matters. Although the response rate after one reminder was modest (45.6%), this is in line with response rates reported following recent surveys of unselected out-of-hours users in the UK[7] or the Netherlands.[28] Some differences were evident between responders and non-responders in respect of age and a proxy for socioeconomic status. Thus, although early indications of generalisability of findings are promising, some caution in interpreting the results is warranted and further investigation justified.

Although we specifically asked users to comment on their most recent contact with the out-of-hours service, there was evidence of some confusion in users' minds regarding which sections of the questionnaire they should complete. Thus, many users reported on the nature of telephone interaction with the healthcare professional, even where this was not the definitive management outcome they had experienced in the course of the clinical contact (as reported by the patient). Even though it may be desirable to link the report of users' experiences of care to a specific interaction with a health professional (eg, for purposes of staff appraisal), it seems unlikely that this can be reliably achieved through a postal questionnaire. In such circumstances, we would advocate the use of exit survey methodology where possible, or by the use of face-to-face or telephone administered questionnaires where clear guidance can be given about the precise element of care in question.

In this study, two of the three areas sampled were known to have extremely low levels of ethnic minority representation (<1% non-white) which probably accounts for the very low ethnic representation among survey respondents (<2.5% overall). Although the third area had higher ethnic mix (approximately 9% non-white) and we did observe a higher ethnic representation in the survey respondents (4.1%), in the absence of reliable information on the ethnicity of service users it is not possible to determine how acceptable the survey is to non-white service user populations.

Validity
Evidence of validity of the instrument was supported by the preliminary qualitative fieldwork undertaken with users, and by the iterative process of questionnaire development. Data obtained provided preliminary evidence of construct validity—two scales were identified within the instrument using exploratory PCA. Although further evaluation of the construct validity of the OPQ is required, users with high levels of overall satisfaction with the care they received had higher scores on both of the scales relating to satisfaction with the out-of-hours consultation, or satisfaction with "entry-access" arrangements.

It was reassuring that, in line with expectations based on the scientific literature,[21 22] older users reported higher levels of satisfaction with the consultation than younger users, as did individuals reporting having received longer consultations compared with those receiving shorter consultations.

Reliability
Although reliability data suggested some item redundancy in the consultation satisfaction scale, we believe that each item provided valuable data on separate aspects of the care provided. The stability of the OPQ was investigated using a test-retest methodology with respondents invited to complete a second questionnaire around 14 days following receipt of their first questionnaire. A rather poor response rate obtained in this part of the study may be due to our failure to adequately alert the recipient that this was a retest questionnaire (perhaps, for example, by the use of an alternative colour) rather than a reminder questionnaire.

Overview
Salisbury and colleagues recently proposed a short instrument[7 28] offering service providers a brief assessment of patient satisfaction with the care provided. Our focus group work[11] suggests that such a brief questionnaire may not address areas which service users feel are important when assessing the quality of out-of-hours primary care services (such as concerns regarding the appropriateness of their requests or the assessment of urgency of the contact by call handlers). The questionnaire developed by van Uden and colleagues,[28] using management specific versions to reduce the length of the schedule incorporated an extended range of items. The

management-specific sampling approach does, however, have limitations. Here we have found evidence of a mismatch between the patient-reported and medical record descriptions of the management received, particularly relating to patients receiving telephone advice. This might not, however, be evidence of inaccurate patient recall and reporting, as the patient may have received subsequent care from the service between the time which they were ''sampled'' and receipt of a questionnaire; a problem which is overcome by using one, longer version capable of capturing experiences across all possible management options.

The ability of providers to adapt service provision as required by, and in the light of their performance against set national standards, is dependent on provision of interpretable data. We believe that some recent questionnaires, although attractive on account of their brevity, may not provide interpretable information in the context of a complex system. A new measure must be flexible enough to capture patients' views around the type of care (or management) they received (eg, telephone advice, home visit, treatment centre visit) and the health professional (doctor, nurse, paramedic) delivering the clinical care. The use of well-designed, shorter and longer questionnaires are both likely to have a place within research and evaluation. Shorter questionnaires can reduce respondent burden and increase response rates, and may be optimal when only a broad overview of service is required, whereas the longer questionnaires can provide a level of detail for a more fine grained analysis of patient experience necessary to support service redesign and development.

CONCLUSION

We have described the preliminary psychometric properties of a questionnaire designed to capture and evaluate users' experiences of out-of-hours care. The questionnaire seems broadly acceptable to users, and appeared reliable and valid in use. Further work is planned to investigate the performance of the instrument against external criteria, and to report on the potential utility of the instrument in informing out-of-hours service planning and configuration.

ACKNOWLEDGEMENTS

. . . .

Authors' affiliations
J L Campbell, A Dickens, S H Richards, P Pound, General Practice and Primary Care, Peninsula Medical School, Exeter, UK
M Greco, Client Focused Evaluation Programme Ltd, The Innovation Centre, University of Exeter, Exeter, UK
P Bower, National Primary Care Research and Development Centre, Manchester, UK

REFERENCES

1 **Starfield B**. *Primary care. Balancing health needs, services, and technology.* New York: Oxford University Press, 1998.
2 **Jones R**. *Oxford textbook of primary medical care.* Oxford: Oxford University Press, 2005.
3 **Department of Health**. *National quality requirements in the delivery of out-of-hours services.* London: Department of Health, 2005.
4 **Department of Health**. *Raising standards for patients, new partnerships in out-of-hours care: an independent review of GP out-of-hours services in England (The Carson Report).* London: Department of Health, 2000.
5 **Department of Health**. *Taking healthcare to the patient: transforming NHS ambulance services (The Bradley Report).* London: Department of Health, 2005.
6 **McKinley RK**, Manku-Scott T, Hastings AM, *et al.* Reliability and validity of a new measure of patient satisfaction with out of hours primary medical care in the United Kingdom: development of a patient questionnaire. *BMJ* 1997;**314**:193–8.
7 **Salisbury C**, Burgess A, Lattimer V, *et al.* Developing a standard short questionnaire for the assessment of patient satisfaction with out-of-hours primary care. *Fam Pract* 2005;**22**:560–9.
8 **van Uden CJ**, Zwietering PJ, Hobma SO, *et al.* Follow-up care by patient's own general practitioner after contact with out-of-hours care. A descriptive study. *BMC Fam Pract* 2005;**6**:23.
9 **Salisbury C**. Postal survey of patients' satisfaction with a general practice out of hours cooperative. *BMJ* 1997;**314**:1594–8.
10 **Coulter A**. Can patients assess the quality of health care? *BMJ* 2006;**333**:1–2.
11 **Campbell JL**, Richards S, Dickens A, *et al.* The Out-of-hours Patient Questionnaire Study; report of study. Exeter: Peninsula Medical School, 2005.
12 **Streiner DL**, Norman GR. *Health measurement scales: a practical guide to their development and use*, 3rd edn. Oxford: Oxford University Press, 2003.
13 **Ramsay J**, Campbell JL, Schroter S, *et al.* The General Practice Assessment Survey (GPAS): tests of data quality and measurement properties. *Fam Pract* 2000;**17**:372–9.
14 **Greco M**, Powell R, Sweeney K. The Improving Practice Questionnaire (IPQ): a practical tool for general practices seeking patient views. *Education for Primary Care* 2003;**14**:440–8.
15 **Greco M**, Cavanagh M, Brownlea A, *et al.* The doctors' interpersonal skills questionnaire (DISQ): a validated instrument for use in GP training. *Education for General Practice* 1999;**10**:256–64.
16 **Salisbury C**. Evaluation of a general practice out of hours cooperative: a questionnaire survey of general practitioners. *BMJ* 1997;**314**:1598–9.
17 **Richards S**, Pound P, Dickens A, *et al.* Exploring users' experiences of accessing out-of-hours primary medical care services. *Qual Saf Health Care* 2007;**16**:469–77.
18 **Townsend P**. Deprivation. *J Soc Policy* 1987;**16**:125–46.
19 **Preacher KJ**, MacCallum RC. Repairing Tom Swift's electric factor analysis machine. *Understanding Statistics* 2003;**2**:13–43.
20 **McHorney CAJ**, Ware JE Jr, Lu JFR, *et al.* The MOS 36 item Short Form Health Survey (SF-36):III. Tests of data quality, scaling assumptions, and reliability across diverse patient groups. *Med Care* 1994;**32**:40–66.
21 **Bowling A**. *Research methods in health: investigating health and health services.* Buckingham: Open University Press, 2002.
22 **Howie JGR**, Porter AM, Heaney DJ, *et al.* Long to short consultation ratio: a proxy measure of quality of care for general practice. *Br J Gen Pract* 1991;**41**:48–54.
23 **Campbell JL**, Ramsay J, Green J. Age, gender, socio-economic, and ethnic differences in patients' assessments of primary health care. *Qual Health Care* 2001;**10**:90–5.
24 **Office of National Statistics**. *General Household Survey, Living in Britain. 2002.* Available at http://www.statistics.gov.uk/ (accessed 1 October 2007).
25 **Cohen J**. *Statistical power analysis for the behavioral sciences.* Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
26 **Gulliford M**, Morgan M, Hughes D, *et al.* Access to health care. Report of a scoping exercise. London: National Co-ordinating Centre for NHS Service Delivery and Organisation Research and Development, 2001.
27 **Nunnally JC**, Bernstein IH. *Psychometric theory*, 3rd edn. New York: McGraw-Hill, 1994.
28 **van Uden CJ**, Ament AJ, Hobma SO, *et al.* Patient satisfaction with out-of-hours primary care in the Netherlands. *BMC Health Serv Res* 2005;**5**:6.