Roy Powell PhD
is Coordinator of the
Peninsula Research and
Development Support Unit
and Research Fellow,
Peninsula College of Medicine
and Dentistry, Noy Scott
House, Royal Devon and
Exeter Foundation Trust.
Helen Powell
is Data Manager, CFEP-UK
Surveys, Innovation Centre,
University of Exeter.
Michael Greco PhD
is Honorary Associate
Professor, School of Medicine,
Griffith University, Australia.

Keywords: patient feedback,
patient surveys, general practice,
service evaluation

The impact of patient survey
feedback in general practice:
The influence of practice size

Roy Powell, Helen Powell and Michael Greco
Received (in revised form): 25th September 2007

Abstract
In two successive years, 2,012 practices surveyed over 400,000
patients per year in the UK using the Improving Practice
Questionnaire (IPQ). In the second survey, of the 27 questionnaire
items, a statistically significant improvement was seen in scores for
12 items, while scores were maintained for 13 items and
significantly decreased for only two items, namely satisfaction with
opening hours and contacting the practice by telephone. Scores for
perceived practitioner capability were maintained regardless of
practice size, showing a continued high level of satisfaction of over
80 per cent in this area. These scores could still be improved,
particularly for considering the patient as a person and taking into
account their personal situation in deciding possible treatments.
Scores for the capacity of the practice to deliver a high-quality
service with regard to supporting services and access were inversely
correlated with practice size. Improvement in these scores indicated
that listening
to patients through the IPQ on more than one occasion can enable
practices to significantly improve patient services, but larger
practices need to work harder at it. Waiting time in the practice was
the lowest scoring item but the efficiency of telephone systems was
also poorly rated. There needs to be greater opportunity for patients
to speak to practitioners on the phone and also to see the
practitioner of their choice, particularly in large practices.

Introduction
The UK NHS is constantly striving towards a more patient-centred
service and a key aspect of this policy is obtaining the views of patients.
Since 1999, the Improving Practice Questionnaire (IPQ) has been a
valuable tool to this end, both for general practices and for health
professionals.1,2

Since April 2004, when the new General Medical Services contract
was introduced, it has been one of the two approved surveys, together
with the General Practice Assessment Questionnaire,3 for the Quality and
Outcomes Framework (QOF) of the contract.4 QOF rewards general
practitioners with points for undertaking measures to improve the quality
of the patient experience. Points are awarded annually for participating in
The impact of patient survey feedback in general practice

an approved survey, reflecting on the results and planning priorities for action with the involvement of patients if possible.

The IPQ is a 27-item, post-consultation, validated survey instrument that combines both the general issues around primary care (domains of access, availability of information, preventative care and health promotion) and assessments of the patient’s experience of the doctor’s interpersonal skills in the medical consultation.\textsuperscript{5,6} The reliability and validity of the IPQ have been established through analysing a large UK dataset of 55,687 patients drawn from 361 practices.\textsuperscript{7} The IPQ assesses two key components of the patient experience. The first component reflects patient perceptions of the clinician’s interpersonal skills and perceived capability, while the second component evaluates capacity issues relating to the accessibility of the practice, the staff attitudes and information giving, and comfort and privacy of the practice facilities.

Few longitudinal studies have been conducted concerning the satisfaction of patients with general practice services. In 2001, however, Greco \textit{et al.} conducted a longitudinal study of the interpersonal skills of GP registrars over four years. This showed how the impact of patient feedback to improve scores over time could be enhanced by education interventions.\textsuperscript{8} In this case, IPQ could be enhanced using critical friends groups.

The aim of this service evaluation was to compare data from practices of various sizes that completed the IPQ on two occasions in two successive annual QOF cycles. Through this evaluation, it was hoped to determine whether practices using the IPQ twice show improved service scores. The null hypotheses being tested were that there would be no change in item scores on the IPQ between the two years, and that there would be no difference between small, medium or large practices.

\section*{Methods}

In total, 2,012 practices in the UK were identified to have carried out the IPQ in both the April 2004 to March 2005 and the April 2005 to March 2006 QOF cycles. This was a large dataset of 822,444 patient responses, consisting of 409,486 questionnaires that were collected in the period April 2004 to March 2005 and 412,958 in the following year. Practices encompassed a large range of list sizes (up to 30,000) in a wide spread of UK locations. In addition to feedback about the surgery, some practices carried out the survey to obtain patient feedback for individual practitioners. Where individual feedback was required, a minimum of 40 completed surveys was necessary to make the data statistically robust.\textsuperscript{9} To meet the requirement of the new General Medical Services contract, 25 patients per 1,000 from the practice’s list size had to be surveyed.

The IPQ is a post-consultation exit survey. Practice staff were advised to hand out questionnaires to consecutive patients so that clinical staff would be unaware which patients were completing them. Patients were requested to complete the questionnaire following their consultation with
the clinician and to rate the items according to their satisfaction with this specific visit. To ensure patient confidentiality and to encourage honest feedback, envelopes for the completed questionnaires were provided and were collected in ballot boxes. Patient anonymity was guaranteed at all times. Questionnaires were processed by an independent company, CFEP-UK Surveys. Data were entered manually into a dedicated database with built-in error checking for verification and were exported into SPSS v.14 for analysis. Questionnaires were excluded if they had fewer than half of the items answered as previous work had shown that these were unreliable and could not be used for calculating component scores.10

The 2,012 practices were divided into tertiles according to their list size rounded to the nearest 100 patients. Thus, small practices were defined as having list sizes with fewer than 4,200 patients, medium-sized practices had between 4,200 and 7,700 patients, and large practices had greater than 7,700 patients. The maximum list size was 29,897 patients. IPQ ‘capacity’ and ‘capability’ scores for the two surveys were analysed using repeated measures analysis of variance comparing the three practice sizes of the two surveys. Individual items were also analysed in this way to give greater detail. Potential confounders were the changes in median age of patients between the two surveys, relative proportions of patient gender, proportion seen by their usual practitioner, and proportions of patients registered less than five years or more than 10 years. No adjustment was made for the number of significance tests performed. However, due to the number of comparisons made, difference was only considered significant if $p \leq 0.001$.

## Results

A small proportion of questionnaires (1,030; 0.13 per cent) were excluded from the analysis because the patients had answered fewer than half of the items on them, making them unreliable. While questionnaire items were completed by most patients, there were some missing demographic data across both surveys. A total of 29,923 patients (3.6 per cent) did not indicate their gender; 68,029 patients (8.3 per cent) did not record whether or not they had seen their usual doctor (or other health professional where appropriate). Some 44,697 patients (5.4 per cent) did not reveal their age, and 22,817 patients (2.8 per cent) did not indicate how long they had been registered with the practice.

The characteristics of the practice population from each period are shown in Table 1. Percentages and averages were calculated from available data on the questionnaires. All demographic data were remarkably similar from both years. These included the age groups, gender, percentage seeing their usual doctor or other health professional and the time that they had been registered with the practice. The majority of patients were female (approximately 66 per cent). The distributions of patients’ ages (see Figure 1) were skewed and bimodal with a trough at age 50. The age distribution was positively skewed for females and negatively skewed for males. Due to this asymmetry, average age was
expressed as medians and interquartile ranges. The most frequent age group was therefore 41–60 years in both surveys, but males were on average 10 years older than females.

The mean item scores for the three practice sizes in each survey are shown in Figures 2–4. Table 2 lists the questions asked in the surveys and gives their mean scores. When ranked in order of their score in the first survey, the top 11 items were all concerned with the health professional’s interpersonal skills and perceived capability. No such items came lower than half-way down the ranking. The lowest scoring items were about the practice itself and its capacity to deliver the service.
The bottom five items were all concerned with access and waiting time issues.

**Capacity scores**

There was a negative correlation between list size and capacity score in both surveys ($r = -0.401, p < 0.001$ and $r = -0.418, p < 0.001$ respectively). Regarding within-practice contrasts, there was a significant increase in capacity score across all three practice size groups at the second survey ($F = 25.276; 1, 2002$ d.f. $p < 0.001$) (see Figure 5). The
The impact of patient survey feedback in general practice

Table 2: IPQ items ordered by question number with their overall rank from the first survey and a comparison of their scores from both surveys derived from repeated measures ANOVA

<table>
<thead>
<tr>
<th>Question</th>
<th>Rank</th>
<th>Mean 2004–05</th>
<th>Mean 2005–06</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1: Your level of satisfaction with the practice’s opening hours</td>
<td>18</td>
<td>3.6894</td>
<td>3.6726</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q2: Ease of contacting the practice on the telephone</td>
<td>23</td>
<td>3.5044</td>
<td>3.5174</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q3: Satisfaction with the day and time arranged for your appointment</td>
<td>17</td>
<td>3.7149</td>
<td>3.7151</td>
<td>0.784</td>
</tr>
<tr>
<td>Q4: Chances of seeing a doctor/nurse within 48/24 hours</td>
<td>24</td>
<td>3.5278</td>
<td>3.5598</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q5: Chances of seeing a doctor/nurse of your choice</td>
<td>25</td>
<td>3.4082</td>
<td>3.4118</td>
<td>0.781</td>
</tr>
<tr>
<td>Q6: Opportunity of speaking to a doctor/nurse on the telephone when necessary</td>
<td>26</td>
<td>3.2884</td>
<td>3.3542</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q7: Comfort level of waiting room (eg chairs, magazines)</td>
<td>21</td>
<td>3.6261</td>
<td>3.6429</td>
<td>0.011</td>
</tr>
<tr>
<td>Q8: Length of time waiting in the practice</td>
<td>27</td>
<td>3.2312</td>
<td>3.2381</td>
<td>0.414</td>
</tr>
<tr>
<td>Q9: My overall satisfaction with this visit to the doctor/nurse</td>
<td>6</td>
<td>4.2149</td>
<td>4.2229</td>
<td>0.152</td>
</tr>
<tr>
<td>Q10: The warmth of the doctor/nurse’s greeting to me</td>
<td>4</td>
<td>4.2468</td>
<td>4.2596</td>
<td>0.007</td>
</tr>
<tr>
<td>Q11: On this visit I would rate the doctor/nurse’s ability to really listen to me</td>
<td>3</td>
<td>4.2586</td>
<td>4.2696</td>
<td>0.029</td>
</tr>
<tr>
<td>Q12: The doctor/nurse’s explanations of things to me</td>
<td>7</td>
<td>4.1866</td>
<td>4.2029</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q13: The extent to which I felt reassured by this doctor/nurse</td>
<td>9</td>
<td>4.1355</td>
<td>4.1518</td>
<td>0.001</td>
</tr>
<tr>
<td>Q14: My confidence in the doctor/nurse’s ability</td>
<td>2</td>
<td>4.2908</td>
<td>4.2989</td>
<td>0.101</td>
</tr>
<tr>
<td>Q15: The opportunity the doctor/nurse gave me to express my concerns or fears</td>
<td>8</td>
<td>4.1794</td>
<td>4.1921</td>
<td>0.005</td>
</tr>
<tr>
<td>Q16: The respect shown to me by this doctor/nurse</td>
<td>1</td>
<td>4.3478</td>
<td>4.3504</td>
<td>0.805</td>
</tr>
<tr>
<td>Q17: The amount of time given to me for this visit</td>
<td>14</td>
<td>3.9215</td>
<td>3.946</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q18: This doctor/nurse’s consideration of my personal situation in deciding a treatment or advising me</td>
<td>11</td>
<td>4.101</td>
<td>4.1113</td>
<td>0.052</td>
</tr>
<tr>
<td>Q19: The doctor/nurse’s concern for me as a person in this visit</td>
<td>10</td>
<td>4.1277</td>
<td>4.143</td>
<td>0.002</td>
</tr>
<tr>
<td>Q20: The recommendation I would give to my friends about this doctor/nurse</td>
<td>5</td>
<td>4.2239</td>
<td>4.234</td>
<td>0.044</td>
</tr>
<tr>
<td>Q21: The manner in which you are treated by the reception staff</td>
<td>12</td>
<td>4.0398</td>
<td>4.0673</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q22: Respect shown for your privacy and confidentiality</td>
<td>13</td>
<td>4.0277</td>
<td>4.0497</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q23: Information provided by the practice about its services (eg repeat prescriptions, test results, cost of private certificates)</td>
<td>15</td>
<td>3.8597</td>
<td>3.8809</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q24: The opportunity for making compliments or complaints to this practice about its service and quality of care</td>
<td>22</td>
<td>3.6007</td>
<td>3.6464</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q25: The information provided by this practice about how to prevent illness and stay healthy (eg alcohol use, health risks of smoking, diet habits, etc)</td>
<td>16</td>
<td>3.7678</td>
<td>3.8025</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q26: The availability and administration of reminder systems for ongoing health checks</td>
<td>20</td>
<td>3.6552</td>
<td>3.6967</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Q27: The practice’s respect of your right to seek a second opinion or complementary medicine</td>
<td>19</td>
<td>3.6663</td>
<td>3.694</td>
<td>&lt;0.001</td>
</tr>
</tbody>
</table>
The mean difference was 0.301 per cent, \( p < 0.001 \). Among the potential confounders, there were significant interactions with the change in patient age (\( F = 44.32; 1,2002 \text{ d.f.} \ p < 0.001 \)) and also the change in proportion that had not seen their usual practitioner (\( F = 49.173; 1,2002 \text{ d.f.} \ p < 0.001 \)). Small practices scored significantly higher than both medium practices (mean difference 4.066 per cent, \( p < 0.001 \)) and large practices (mean difference 6.136 per cent, \( p < 0.001 \)). Furthermore, medium practices scored higher than large practices (mean difference 2.07, \( p < 0.001 \)).

Most of the capacity items increased in score across all practice size groups except for two. These were Q1 ‘satisfaction with the practice’s opening hours’, which decreased for medium sized and large practices, and Q2 ‘ease of contacting the practice on the telephone’, which decreased across all three practice size groups (\( p < 0.001 \) in both cases).

**Capability scores**

There was no correlation between list size and capability score in either survey. Within-practice contrasts indicated a significant increase in capability score at the second survey (\( F = 7.184; 1,2002 \text{ d.f.} \ p = 0.007 \)) (see Figure 6). The mean difference was 0.148 per cent. There were significant interactions with three confounders. These were the change in patient age (\( F = 31.867; 1,2002 \text{ d.f.} \ p < 0.001 \)), the change in proportion that had not seen their usual practitioner (\( F = 33.227; 1,2002 \text{ d.f.} \ p < 0.001 \)) and the change in the proportion registered with the practice for more than 10 years (\( F = 13.481; 1,2002 \text{ d.f.} \ p < 0.001 \)). However, there were no significant differences between the three practice sizes.
sizes. Marginal mean values for small, medium and large practices were 83.846 per cent, 83.738 per cent and 83.964 per cent respectively. Across this narrow range, however, the score for small practices changed the most, starting below that for medium and large practices and actually overtaking them in the second survey (see Figure 6). In fact, improvements were seen in all the capability questions among small practices. This was not always the case for medium and large practices (see Figure 3).

Allowing for confounders, statistically significant improvements were seen in 12 items on the IPQ at the 0.1 per cent level (i.e., $p \leq 0.001$). Two items received a significant decrease in score. These were Q1 ‘level of satisfaction with the practice’s opening hours’ and Q2 ‘ease of contacting the practice on the telephone’ (both $p < 0.001$). Three items in the health professional’s capability domain improved statistically significantly: Q12 ‘doctor/nurse’s explanations of things to me’ ($p < 0.001$), Q13 ‘Extent to which I felt reassured by this doctor/nurse’ ($p = 0.001$) and Q17 ‘amount of time given to me on this visit’ ($p < 0.001$). However, scores for clinicians’ interpersonal skills were already relatively high. The remaining nine items that improved significantly were from other areas of the questionnaire, which dealt with the practice’s capacity to deliver the service. The largest improvement of all was in Q6 ‘opportunity of speaking to the doctor on the telephone when necessary’ ($p < 0.001$), which also had the second lowest starting value of all the items on the questionnaire (after Q8 ‘length of time waiting in the practice’). Eight other items that showed significant improvement in terms of their ranked increase in scores were all ‘capacity’ items (Q4 and Q21–Q27 inclusive).
Discussion

This study appears to show that reflecting and acting on patient feedback from the IPQ positively influenced practice scores when the survey was carried out again 12 months later. The demographic features of the two surveys were remarkably similar. The major differences were that in the second survey the proportion of patients over 60 years old had increased by 0.7 per cent. Men were on average one year older and 0.4 per cent fewer had been registered for less than five years. IPQ capacity and capability scores increased in the second survey even when allowing for these confounders. There was a negative correlation between list size and capacity score but no such relationship with capability score. Small practices therefore had high capacity ratings. Even so, there were improvements in capacity score for all three practice size groups. Statistically significant improvements were made in 12 of the 27 IPQ items in the second survey and most of these related to aspects of the practice’s capacity to deliver a high-quality service rather than focusing on the capability of the health professionals, which already received relatively high scores. There was a significant increase in capability score but practice size made no difference to it. Among individual items, the highest scoring ones related to practitioners’ interpersonal skills and perceived capability. The lowest scoring ones were about access and waiting times.

Small practices improved in every item except Q2 ‘ease of contacting the practice on the telephone’. For medium and large practices, one additional item significantly decreased in score: Q1 ‘level of satisfaction with the practice’s opening hours’. These related to a period covered by the surveys when GP services had been limited to 08.00–18.30 on normal working days and many Saturday surgeries had been curtailed. While accessing the practice had not improved (opening hours and telephone access), accessing the practitioners themselves does appear to have improved (opportunity of speaking to a practitioner on the phone, chances of seeing a doctor/nurse within 48/24 hours and amount of time given by the practitioner) and is reflected in the scores.

The results of this evaluation underline the continued importance to patients of the respect shown to them by doctors and other health professionals, their confidence in the professional’s clinical ability and how much they felt they listened to them. The importance of the two-year comparison is that it emphasises patients’ appreciation of improvements that can be made. These reflected areas of the practice where it is perhaps easier for changes to be initiated, such as the respect shown to them by reception staff, the accessibility of the practice, its systems for reminders and providing prevention information, the availability of the doctor or other health professional on the phone and the amount of time that they are able to spend with them. Patients have also valued improvement in the opportunity to give practices feedback and the recognition of their right to a second opinion. In a randomised controlled trial of critical friends groups it was shown that provision of facilities to allow greater privacy, for example, were easy to achieve and
were valued by patients as observed in focus group interviews and second IPQ survey scores.\textsuperscript{11}

One weakness of the study is that the way in which questionnaires are distributed precludes the possibility of matching them year on year at the individual patient level. On the other hand, the system means that a high response rate is guaranteed. The best that can be achieved is to match on practices. There may also be different numbers of returned surveys from participating clinicians on each occasion.

At the time when this analysis was done, no data were available on the method of reflection or future planning that practices carried out after receiving the results of their survey. Different approaches may have led to variations in the amount of improvement seen in the subsequent survey. Another study has found that practices which discussed their survey results with patient (or critical friends) groups were more likely to improve their scores than those practices that did not.\textsuperscript{12}

A strength of the study was that, as in previous work with the IPQ, the questionnaires were well completed by patients, with few item response problems.\textsuperscript{13}

It is known from previous work\textsuperscript{14} that the demographic details recorded in the IPQ are potential confounders. Generally, men give higher scores than women and older people give higher scores than young people. Furthermore, lower scores are given if patients do not see their usual practitioner. In addition, patients who have been registered with the practice for longer (more than 10 years) will give higher scores than those registered only a short time (less than five years). In this study one could therefore expect that the slightly larger proportion of older people (especially men) in the second survey will tend to increase overall scores. Additionally, the slight reduction in the proportion of female patients will also tend to increase scores. By building these confounders into the models for capacity and capability, it was possible to get a more realistic idea of the effect of undergoing the survey for a second time. It was found that, for most of the IPQ items, the change in patients’ average age and the change in proportion seeing their own doctor were significant confounders.

The study also bears out previous findings from IPQ surveys, particularly the ranking of items in terms of their scores. Whenever the IPQ has been conducted in the UK in the past, the highest scoring items are invariably the ones relating to the interpersonal skills and capability of the doctor or other clinician that the patient has seen on that visit, and the lowest are practice services.\textsuperscript{15} This fact has been referred to as the ‘tomato sauce’ analogy. Patients like the product — that is the consultation with the clinician — but have trouble getting it out of the bottle or accessing it.

The highest scoring item in this study was Q16 ‘the respect shown to me by this doctor’. The consistently lowest scoring item in the capability section was Q17 ‘length of time given for the consultation’. Likewise, the two lowest scoring items of the survey were ‘chances of speaking to the doctor on the telephone when necessary’ and ‘length of
time waiting in the practice to see the doctor’. In fact, the worst scoring items were mostly about access (Q2, Q4–Q6 and Q8 in this survey). The first four of these are about access to the practice or the doctor or other health professional. These findings are all consistent with previous work.16–18

Overall, it has been shown that the IPQ is a useful tool to capture patients’ views on all aspects of general practice. Giving clinicians and practice staff a snapshot of their patients’ opinions on their capability or service allows areas of excellence, or areas where improvement is possible, to be clearly pinpointed. This can then allow positive changes to occur to improve the patient experience.19 Further work is required to see if the ability to improve in all these areas varies with practice size. It would also be useful to see if there are socioeconomic or geographical variations.

To look at the ways in which practices are reflecting and acting on their IPQ results, a more qualitative study is currently being carried out to ascertain identified priorities and tangible changes made to practices. Simply doing a survey may not improve performance;20 however, this large evaluation seems to indicate that practice scores improve when targeted quality improvement activities are undertaken.

Conclusion
Feedback from over 800,000 patients from practices carrying out the IPQ twice has shown continued high levels of satisfaction with general practice clinical services (capability) and improved satisfaction with supporting services and access (capacity). However, the latter were inversely correlated with practice size but also improved regardless of size.

Learning points from this study include the fact that all practices can improve, especially in areas of their capacity, but large practices start from a lower baseline and will need to work harder at it. Large practices scored significantly worse than small practices on every capacity item. They scored worse than medium-sized practices for all items except Q1 ‘level of satisfaction with the practice’s opening hours’, Q4 ‘chances of seeing a doctor/nurse within 48/24 hours’, Q7 ‘comfort level of waiting room (eg chairs, magazines)’ and Q8 ‘length of time waiting in the practice’. In fact, patients gave this waiting time the worst scores of all. Practices need to consider the efficiency of their telephone systems too because their scores worsened and were particularly low in larger practices. They need to allow greater opportunity for patients to speak to practitioners on the phone and also to see the practitioner of their choice, particularly in large practices. Perceived practitioner capability was highly rated regardless of practice size but could still be improved, particularly with regard to considering the patient as a person and taking into account their personal situation in deciding possible treatments for them. Improvement in scores indicated that listening to patients through the IPQ on more than one occasion can enable practices to improve patient services significantly, but this may be more difficult for larger practices.
Acknowledgments

The authors would like to thank CFEP-UK Surveys staff for their support, and the practices and patients who participated in the study.

References


5. Greco et al., ref. 1 above.

6. Greco and Carter, ref. 2 above.

7. Greco et al., ref. 1 above.


10. Greco et al., ref. 1 above.

11. Greco et al., ref. 9 above.


13. Greco et al., ref. 1 above.

14. Ibid.

15. Ibid.

16. Ibid.

17. Ware and Hays, ref. 9 above.


19. Ibid.