Bias in patient assessments of general practice: General Practice Assessment Survey scores in surgery and postal responders

Peter Bower and Martin O Roland

SUMMARY
Patient-based measures of the quality of primary care are increasingly important. However, their effective use requires bias to be minimised. Scores on the General Practice Assessment Survey (GPAS) differ according to whether patients are surveyed in the surgery or by post. It is not clear whether these differences relate to the mode of administration or to the types of patients who complete the scale in postal and surgery samples. Regression indicates that the bias reflects both effects and should be considered when GPAS scores are being interpreted.

Keywords: patient satisfaction; quality assessment; selection bias.

Introduction
PATIENT assessments of primary care are important for quality improvement, and the popularity of questionnaires completed by patients has led to interest in aspects of questionnaire design and administration and their relationships to bias, i.e. systematic under- or over-estimation.

One popular questionnaire is the General Practice Assessment Survey (GPAS). The GPAS has been used in surveys of surgery attenders and in postal surveys of practice populations. Scores in surgery responders are about 5% higher overall, but rise to 10% on the receptionist scale, which may have important implications for the interpretation of data.

These differences may relate to task effects (differences related to the context in which the questionnaires are completed; for example, by postal and surgery responders), or to selection effects (relating to differences in the sociodemographic nature of surgery and postal samples). Age, ethnicity, and socioeconomic status are associated with GPAS scores.

This study examined these two hypotheses by determining whether the scores of surgery and postal responders differ when sociodemographic differences are statistically controlled.

Method
The full GPAS dataset included 21,905 patients, 56.8% of whom were responders from surgeries. Data from postal surveys involved random samples from practice lists, while surgery samples involved questionnaire administration by practice staff to consecutive surgery attenders. Only those GPAS scale items present in versions 1 and 2, i.e. access, receptionists, continuity, communication, personal care, knowledge of the patient, nursing, and general satisfaction, were analysed (Figure 1). All except one of the GPAS assessment items use a six-point response format, ranging from ‘very poor’ to ‘excellent’. For reporting purposes scale scores (between zero and 100) are calculated as a percentage of the maximum possible score.

The following sociodemographic characteristics were used in the analysis: chronic illness (‘yes’ or ‘no’); sex; age; ethnic minority (‘white’ or ‘other’); and accommodation (‘owner-occupier’ or ‘other’). Responders under the age of 16 years were excluded. Patients for whom there was missing data on any of the study variables (the GPAS scales or the demographic variables) were excluded, to provide a...
consistent sample size for analyses. However, because a significant proportion of patients had not consulted a nurse, a reduced dataset was used for the analysis of the GPAS nursing scale.

Analysis
To determine the effect of surgery and postal administration, multiple regression (using SPSS version 9) was used to examine the relationship between type of administration and GPAS scale scores, controlling for sociodemographic differences. All variables were entered into the regression simultaneously. The unstandardised regression coefficient relating to the binary administration variable (i.e. surgery or postal) estimated the difference in scores on each GPAS scale between surgery and postal responders, holding all other demographic factors constant.7 GPAS scale scores are generally negatively skewed, which can influence multiple regression. Although transformations were attempted to reduce skew, these had little substantive effect on the results.
Results

The included sample size was 13,534, 61.8% of the total dataset, with 56.6% of surgery responders and 68.6% of postal responders included. Included and excluded patients were compared on demographic characteristics where these were recorded (n = 16,223). Comparisons indicated that patients who were included were more likely to be white (86.4% versus 81.4%), owner-occupiers (63.9% versus 57.3%), and chronically ill (37.6% versus 33.4%), but did not differ in sex (62.8% female versus 61.2%) or age (mean age = 46.8 years in both groups).

Analysis of demographic differences between postal and surgery responders indicated that surgery responders were more likely to be younger (mean age = 44.3 years versus 49.5 years), female (65.0% versus 60.4%), from an ethnic minority (20.6% versus 6.0%), to have a chronic illness (44.4% versus 30.2%) and less likely to be owner-occupiers (54.4% versus 74.2%).

Table 1 shows the mean scores for each GPAS scale, the unadjusted difference between postal and surgery responders, and the difference derived from the multiple regression (i.e. the unstandardised regression coefficient), controlling for demographic factors. In all cases, the adjusted difference was greater than the original difference.

Discussion

GPAS scores are higher (i.e. more positive) in surgery than postal responders. The data suggests that the differences between surgery and postal responders relate both to a response set issue and to a selection effect. Generally, administering GPAS in the surgery accesses patients who are more likely to be in those groups associated with lower GPAS scores, i.e. from an ethnic minority, less affluent, and younger.3 When these effects are held constant, the bias towards higher scores in surgery responders is increased.

This study cannot determine the cause of the bias. Given that the bias is largest in the receptionist scale, and that reception staff are often involved in distributing questionnaires,2 patient concerns about confidentiality may result in socially desirable responses (i.e. more positive assessments), similar to the bias that is potentially found in face-to-face interviews.1

The substantive significance of these differences is unclear. According to the GPAS manual,5 practices should be highlighted if they have scores 10 points above or below the mean in the locality. Only the bias in the receptionist scale approached this magnitude. Nevertheless, any bias may be important if it is in addition to modest (though real) differences in patient assessments that result in a practice being identified as an outlier.

It should be noted that the current results do not necessarily suggest that postal administration is superior in all cases. However, researchers using patient assessments will need to be aware of the biases associated with differences in administration when interpreting data.

References

6. www.gpas.co.uk

Acknowledgements

The authors would like to thank all those who provided data for the analyses, especially Stephen Campbell of the NPCRDC, and John Campbell, from the Department of General Practice and Primary Care, Guy’s, King’s and St Thomas’ School of Medicine.