

Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England

Martin Roland, professor of health services research,¹ Marc Elliott, senior statistician,² Georgios Lyratzopoulos, clinical senior research associate,¹ Josephine Barbieri, research assistant,¹ Richard A Parker, research assistant,¹ Patten Smith, director of research methods,³ Peter Bower, reader in health services research,⁴ John Campbell, professor of general practice and primary care⁵

¹Department of Public Health and Primary Care, University of Cambridge, Institute of Public Health, Cambridge CB2 0SR

²RAND Corporation, 1776 Main Street, Santa Monica, CA 90401-3208, USA

³Ipsos MORI, London SE1 1FY

⁴National Primary Care Research and Development Centre, University of Manchester, Manchester M13 9PL

⁵Peninsula Medical School, Smeall Building, Exeter EX1 2LU

Correspondence to: M Roland
mr108@cam.ac.uk

Cite this as: *BMJ* 2009;339:b3851
doi:10.1136/bmj.b3851

ABSTRACT

Objective To assess the robustness of patient responses to a new national survey of patient experience as a basis for providing financial incentives to doctors.

Design Analysis of the representativeness of the respondents to the GP Patient Survey compared with those who were sampled (5.5 million patients registered with 8273 general practices in England in January 2009) and with the general population. Analysis of non-response bias looked at the relation between practice response rates and scores on the survey. Analysis of the reliability of the survey estimated the proportion of the variance of practice scores attributable to true differences between practices.

Results The overall response rate was 38.2% (2.2 million responses), which is comparable to that in surveys using similar methodology in the UK. Men, young adults, and people living in deprived areas were under-represented among respondents. However, for questions related to pay for performance, there was no systematic association between response rates and questionnaire scores. Two questions which triggered payments to general practitioners were reliable measures of practice performance, with average practice-level reliability coefficients of 93.2% and 95.0%. Less than 3% and 0.5% of practices had fewer than the number of responses required to achieve conventional reliability levels of 90% and 70%. A change to the payment formula in 2009 resulted in an increase in the average impact of random variation in patient scores on payments to general practitioners compared with payments made in 2007 and 2008.

Conclusions There is little evidence to support the concern of some general practitioners that low response rates and selective non-response bias have led to systematic unfairness in payments attached to questionnaire scores. The study raises issues relating to the validity and reliability of payments based on patient surveys and provides lessons for the UK and for other countries considering the use of patient experience as part of pay for performance schemes.

INTRODUCTION

Financial incentives to improve quality of care, sometimes called pay for performance schemes, have been introduced recently in many countries, including the United States,^{1,2} Spain,³ and Australia.⁴ The United Kingdom embarked on the most ambitious of these schemes in 2004 with an initiative in which 25% of general practitioners' pay was tied to a complex set of quality indicators, the quality and outcomes framework.⁵ In common with other countries, most of the indicators in the original UK framework related to clinical care. These incentives were associated with accelerated improvement for some aspects of chronic disease management⁶ and a reduction in inequalities in the delivery of primary care.⁷

Patient experience is an additional important component of quality, and questionnaires to measure patient experience have been widely used to assess care.⁸⁻¹³ Some schemes have included patient experience as an incentivised element of quality, either with providers incentivised to report on patient experience¹⁴ or with a direct link between patients' evaluation and doctors' income.^{15,16} In 2007 and 2008, financial incentives were attached to the results of a survey in the UK in which patients reported on how easy they found it to get an appointment with their doctor. In 2009, as part of negotiations between NHS Employers and the General Practitioners Committee of the British Medical Association, this incentive was included in the quality and outcomes framework: £68m (\$109m) of general practitioners' pay was tied to patients' reported experience of access to care (getting urgent appointments and being able to book ahead).¹⁷

Following the development of a substantially extended survey instrument during the second half of 2008, the new GP Patient Survey was administered by Ipsos MORI between January and March 2009. Postal questionnaires in English were sent with reply paid envelopes to 5 660 232 patients aged 18 or over who had been continuously registered with a general practice in England for at least six months. Non-

respondents were sent reminders and additional questionnaires one month and two months after the initial mailing unless they indicated that they wished to opt out of the survey. Questionnaires could be completed in 13 non-English languages on line or by telephone, and in British Sign Language on line (3.4%, 0.01%, and 0.03% of responses respectively). Patients could also complete the survey in Braille or large print, but these options produced no responses. Out of 8426 practices in England, 135 with fewer than 50 qualifying patients at the time of sampling were excluded, as were another 18 practices, leaving 8273 (mean list size 4946 patients) contributing to the data. To get sufficient precision on responses related to payments, stratified random samples were drawn from practice lists, which resulted in an average of 260 responses from practices of all sizes. The sampling method involved over-sampling from small practices and from practices with low response rates in previous surveys conducted by Ipsos MORI. The mean number of patients invited to take part per practice was 684. Details of the development of the survey—including piloting, cognitive interviewing, sampling, sample size calculations, and details of response rates—have been reported elsewhere.^{18 19}

Results of the survey were published in July 2009 for each general practice in England at www.gp-patient.co.uk/results. Patients generally reported positive experiences and high levels of satisfaction with their general practices. Questions where difficulty or dissatisfaction with care was expressed by more than 15% of patients included privacy in reception, getting through on the telephone, speaking to a doctor on the phone, being able to book appointments ahead, waiting to be seen by a doctor, and being able to see a doctor of the patient's choice. A change in the payment formula that related questionnaire scores to general practitioner income (see appendix 1 on bmj.com) meant that many general practitioners did not receive the income they expected from their survey results. Political discussion focused on the validity and reliability of the survey, and many appeals were mounted by general practitioners in relation to the payments they had received. These appeals revolved around two core issues—that insufficient patients were surveyed in some practices, and that the poor response rate meant that the survey results were biased.

Using patient experience to measure quality raises issues of reliability and validity that are not encountered with more commonly used measures of clinical quality. Reliability issues include the precision of measurement and the ability of questions to discriminate between practices.²⁰ In terms of validity, there are several reasons why responses to patient questionnaires need to be interpreted with caution, including bias associated with selective non-response,²¹ inaccuracy in recall,²² the effect of context on patient response,²³ different expectations in different population groups,²⁴ and other sources of evidence on which patients draw when making judgments.^{25 26} In addition, there are few truly independent methods of assessing external

validity. For example, compared with the US,²⁷ UK patients rarely change practice because they are dissatisfied with their care, so voluntary movement between practices cannot readily be used as an external measure of validity.

As part of an ongoing programme of research on patient experience, we explored issues of reliability and non-response bias in the UK survey, outlining important lessons for the UK and for other countries wishing to link physician payments to the results of patient questionnaires.

METHOD

We examined the representativeness of survey respondents in two ways. Firstly, we compared the patients who responded to the survey with those who were sampled in term of age, sex, and deprivation. Secondly, we compared the respondents with estimates from the general population in terms of age, sex, deprivation, and ethnicity.

We then examined the relation between practice response rates and the scores obtained on the two questions that triggered payments to general practitioners, directly and controlling for patient characteristics. This was done to determine whether practices with lower response rates might have received responses from patients with systematically more positive or negative experiences, leading to biased scores.

Finally, we calculated the reliability of practice-level mean scores for the two questions that related to payments to general practitioners. The reliability coefficient (an index of 0 to 1) represents the proportion of the variance of practice-level mean scores that is attributable to true differences between practices, as opposed to differences which might be attributable to incomplete measurement. Reliability greater than 70% indicates acceptable reliability; reliabilities of 80-90% are regarded as preferable for higher stakes applications such as pay for performance.²⁸

Details of the statistical methods used for these analyses are shown in appendix 2 on bmj.com. Analyses were carried out using SPSS version 15 for descriptive analyses, and STATA version 10 for regression models, weighting analyses, and analyses of reliability and non-response bias.

RESULTS

The overall response rate was 38.2% (2 163 456), which is comparable to that achieved in surveys using similar methodology in the UK, but lower than the 44% and 41% achieved in the shorter 2007 and 2008 surveys which measured only patients' experience of access to primary care.^{29 30} In part, the response rate reflects deliberate over-sampling from practices with historically low response rates to obtain sufficient responses for each practice. The overall response rate weighted to adjust for this aspect of the survey design was 42.3%. An average of 261.7 (standard deviation 39.9) patients responded in each practice; 361 practices (4.4%) had

fewer than 200 responses, and 47 practices (0.6%) had fewer than 50 responses.

Representativeness of patients who responded to the survey

Table 1 shows the sociodemographic factors that predicted survey response. The youngest patients (age 18-29 years) were the least likely to respond. The odds of responding increased slightly in middle aged patients, then increased substantially in elderly patients to a peak at the ages 70-79, where the odds of responding were 5.54 times as high as for the reference group aged 18-29. Response propensity then declined beyond age 79, with patients aged 80-89 (odds ratio 3.59) less likely to respond than those aged 60-69 (odds ratio 4.53). For socioeconomic deprivation, the odds of responding declined approximately linearly with increasing deprivation, with those in the second deprivation fifth having 10% lower odds of response than those in the reference top fifth (odds ratio 0.90 (95% CI 0.90 to 0.91)) and those in the bottom fifth of deprivation having about half the odds of response as those in the top fifth (odds ratio 0.52 (0.52 to 0.53)). For sex, the odds of responding were 41% lower for men than for women, allowing for the effects of age and deprivation.

Table 2 shows the proportions of respondents from different age and deprivation bands and ethnic groups compared with recent population based estimates for England from the Office for National Statistics. These are shown in the table in terms of weighted

Table 1 Predictors of patient response to GP Patient Survey 2009 by age, sex, and deprivation (n=5658740*)

Factor	Odds ratio	Standard error (95% CI)
Age (years):		
18-29†	1.00	
30-39	1.06	0.003 (1.06 to 1.07)
40-49	1.48	0.004 (1.48 to 1.49)
50-59	2.54	0.007 (2.52 to 2.55)
60-69	4.53	0.015 (4.50 to 4.56)
70-79	5.54	0.026 (5.49 to 5.59)
80-89	3.59	0.017 (3.56 to 3.62)
≥90	1.46	0.016 (1.43 to 1.49)
Sex:		
Female†	1.00	
Male	0.59	0.001 (0.59 to 0.59)
Socioeconomic deprivation group‡:		
Group 1 (least deprived)†	1.00	
Group 2	0.90	0.001 (0.90 to 0.91)
Group 3	0.79	0.003 (0.78 to 0.79)
Group 4	0.64	0.002 (0.64 to 0.65)
Group 5 (most deprived)	0.52	0.002 (0.52 to 0.53)

*Excludes 1492 cases with missing value (1477 with missing deprivation group and 15 with missing sex).

†Reference category.

‡Based on individual patient postcodes and using population based Lower Super-Output Area Indices of Multiple Deprivation for 2007 with cut points for five equal groups of 8.257, 13.525, 20.741, and 33.511.

characteristics of respondents, but analyses using unweighted estimates showed a similar overall pattern.

These data show under-representation of men, younger people, and people from deprived areas among survey respondents. For ethnicity, the proportion of respondents describing themselves as white British was similar to the general population. There was under-representation of respondents describing their ethnicity as “Asian/Asian British,” “black/black British,” “mixed,” and “Chinese,” with 2.7% fewer respondents from these four ethnic groups than would have been expected from the general population. At the same time, there was a substantial over-representation of individuals describing their ethnicity as “other” (3% more than the proportion in the general population).

Non-response bias

Practices with high rates of response had slightly higher scores on the two access questions relating to general practitioner payments (PE7 and PE8, see details in appendix 1). Pearson (Spearman) correlations were 0.34 (0.34) for PE7 and 0.15 (0.18) for PE8 (P<0.001 for each). However, these associations were almost

Table 2 Comparison of weighted demographic characteristics of respondents to GP Patient Survey 2009 with estimates for the resident population of England aged ≥ 18 years (mid-year 2007). All values are percentages

Characteristic	Survey respondents after weighting*	England population estimates
Sex:		
Male	42.2	48.6
Age (years):		
18-29	10.4	20.4
30-39	13.2	17.8
40-49	17.1	18.7
50-59	18.4	15.5
60-69	19.7	12.8
70-79	14.1	9.0
80-89	6.4	4.9
≥90	0.7	0.9
Socioeconomic deprivation group‡:		
Group 1 (least deprived)	21.1	20.0
Group 2	20.5	20.0
Group 3	20.0	20.0
Group 4	19.4	20.0
Group 5 (most deprived)	19.0	20.0
Ethnicity‡:		
White British	88.9	89.3
Mixed	0.6	1.1
Asian/Asian British	3.9	5.3
Black/black British	2.2	2.7
Chinese	0.5	0.8
Other	3.8	0.8

*Practice-level weights after stratification, calculated as the ratio of eligible individuals to respondents within each practice.

†Based on estimates as in table 1.

‡Based on the 97.0% of respondents providing ethnicity information.

entirely a function of the demographic characteristics of practices, as the partial correlations of response rate with the two payment questions were weak and inconsistent in direction once we controlled for demographic factors (0.04 and -0.07 for PE7 and PE8 respectively, $P < 0.05$ for each). The response rate on its own was associated with less than 0.2% of the variance in both PE7 and PE8, and with practice PE7 scores increasing by 0.4% and PE8 scores decreasing by 1.1% for a 10% increase in practice response rate.

Practice-level reliability

The two questions on access to care which related directly to payments to general practitioners (PE7 and PE8) had intra-class correlations of 0.079 and 0.125 respectively, resulting in practice-level reliability coefficients of 93.2% and 95% at the overall mean number of 262 responses per practice (with a mean number of 160 and 134 responses for the two payment questions). In total, 97.0% and 96.9% of practices achieved the 90% threshold for reliability for PE7 and PE8, and 99.5% of practices achieved the number of responses required for 70% reliability. To meet the higher standard of 90% reliability, 105 and 64 responses to the two payment questions were needed.

Although the survey met internationally recognised standards for high reliability,²⁹ a change in the payment formula between 2008 and 2009 increased the effect of random variation on payments for many practices (see appendix 1).

DISCUSSION

The quality and outcomes framework has focused British general practitioners' attention on clinical aspects of care,^{31,32} and a focus of incentives on more holistic aspects of care, including patient experience, may seem a logical extension of the scheme. As with clinical indicators, patient experience was measured at the level of the practice (averaging between four and five doctors per practice in the UK) even though this may itself obscure a substantial amount of variation that is likely to occur between individual physicians.³³

The 2009 GP Patient Survey had an overall response rate of 38.2%, which is 3-6% lower than the other recent access surveys in primary care, though the rate in part reflects deliberate over-sampling from practices known to have low response rates. Many general practitioners expressed concern that these low rates of response would lead to unreliable scores, with particular concern expressed about the questions which related to payment. Previous research suggests that, as long as rigorous probability sampling processes are followed, the association between response rates and non-response bias is weak.³⁴⁻³⁶ In surveys of healthcare experiences, it is generally the case that, after controlling for patient demographic characteristics, patients with more positive experiences are more likely to respond than those with less positive experiences,³⁷ and there has been some concern that low response rates may disproportionately omit those patients with

poor experiences. This would upwardly bias the scores of providers with low response rates relative to their peers with higher response rates, and some have proposed corrections in attempts to address this possibility.³⁷ For the UK GP Patient Survey, we found little evidence that variation in response rates would result in any systematic disadvantage to practices with either low or high rates of response for the questions that were associated with payments to general practitioners.

We observed patterns of non-response that are similar to those in other surveys³⁷⁻³⁹—with men, young people, and people living in deprived areas being less likely to respond to the questionnaire. In some ways, this reflects consulting experience, since women and older people are more likely to consult a doctor. However, the low response rate from deprived areas reinforces systematic under-representation of the views of people living in deprived areas as they are on average high users of general practice services.

The sampling procedure produced highly reliable estimates of patient experience at the average number of 262 completed surveys per practice. For the 3% of practices with fewer than the 105 and 64 responses to the two payment questions needed to meet the highest standards of 90% reliability (for PE7 and PE8 respectively), we recommend that larger numbers of patients from these practices should be sampled in future rounds of the survey. The overall effect of random variation on payments was increased compared with previous years because of a change in the payment formula, and we discuss this in more detail in the appendices to this paper.

Information on patient experience from the GP Patient Survey is now publicly available for English practices. However, practitioners may not engage with these new measures of quality if they believe them to be flawed in terms of validity or reliability. The analyses here suggest that the current survey procedures result in reliable survey estimates of performance at the practice level on the pay for performance items which we examined. However, the attachment of large amounts of payment may also have led practitioners to focus excessively on the some technical aspects of the survey, and limit their value as a quality improvement tool. Future research should focus on the validity of the responses in addition to establishing whether case mix differences in practice populations affect the equivalence of measures of patient experience.

The UK experience of integrating survey based measures of patient experience into pay for performance initiatives has several lessons for other countries. Firstly, such measures must be designed to produce reliable and valid estimates of performance at the intended level of measurement (such as practice or practitioner). Secondly, financial incentives must be designed so that true differences between practices are large in relation to background random variation. Reliable survey measurement is a necessary but not sufficient condition for robust pay for performance

WHAT IS ALREADY KNOWN ON THIS TOPIC

Patient experience is an important component of quality of health care

Some countries are introducing patient experience as part of pay for performance schemes for healthcare workers

UK general practices are now paid on the basis of patient responses to two questions on getting appointments which are part of a larger national survey

WHAT THIS STUDY ADDS

The questions associated with payment to general practitioners showed a high degree of reliability

Practices with low rates of response were neither advantaged nor disadvantaged in terms of patient scores or payments received

A change to the payment formula in 2009 increased the impact of random variation on practice payments

based on patient experience. Designers of payment formulas typically pay close attention to the incentives they intend to create, but may not consider the ways in which the formula can erode the reliability of an otherwise valid measure. Pay for performance initiatives are unlikely to be effective unless there is a strong correspondence between actual performance and compensation. Thirdly, few doctors are likely to be familiar with technical aspects of survey methodology, and the linkage of large payments to surveys may create considerable disquiet. In one US example, public reporting of patient experience data for hospital care preceded an anticipated pay for performance scheme, providing an opportunity to address these concerns before the addition of significant financial stakes.⁴⁰

Patient reported measures of quality are an important aspect of care, and the GP Patient Survey represents a major opportunity to improve care on a national scale. None the less, additional refinements of the measurement or compensation process and ongoing dialogue with practising doctors will be essential if the survey is to play an important role in improving patient experience in the UK.

We acknowledge the assistance from a wider team involved in development of the survey at Ipsos MORI, including Becca Maeso, Helen Rowley, Juliet Brown, Lisa Valade-DeMelo, and Sonja Nissen. We are also grateful to Guy Watkins for his description of the GP payment system. Contributors. All authors were involved in designing the analyses, reviewing intermediate analyses, and contributing to the final paper. MR is the guarantor of the paper.

Funding. The study was funded with a grant from the Department of Health. The opinions expressed are those of the authors and not of the department. All authors had full access to the data and take responsibility for the accuracy of the data analysis. The Department of Health was not involved in the analysis of the data.

Competing interests: PS is an employee of Ipsos MORI, which developed and delivered the GP Patient Survey for the Department of Health. MR and JC act as academic advisers to Ipsos MORI for the survey.

Ethical approval: Not required.

Data sharing: The results of the GP Patient Survey are publicly available at www.gp-patient.co.uk/results

- Rosenthal MB, Frank RG, Li Z, Epstein AM. Early experience with pay-for-performance: from concept to practice. *JAMA* 2005;294:1788-93.

- Rosenthal MB, Landon BE, Normand SL, Frank RG, Epstein AM. Pay for performance in commercial HMOs. *N Engl J Med* 2006;355:1895-902.
- Benavent J, Juan C, Clos J, Sequeira E, Gimferrer N, Vilaseca J. Using pay-for-performance to introduce changes in primary healthcare centres in Spain: first year results. *Qual Primary Care* 2009;17:123-31.
- Scott IA. Pay for performance in health care: strategic issues for Australian experiments. *Med J Aust* 2007;187:31-5.
- Roland M. Linking physician pay to quality of care: a major experiment in the UK. *N Engl J Med* 2004;351:1448-54.
- Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. Effects of pay-for-performance on the quality of primary care in England. *N Engl J Med* 2009;361:368-78.
- Doran T, Fullwood C, Kontopantelis E, Reeves D. Effect of financial incentives on inequalities in the delivery of primary clinical care in England: analysis of clinical activity indicators for the quality and outcomes framework. *Lancet* 2009;372:728-36.
- Greco M, Sweeney K, Brownlea A, McGovern J. The practice accreditation and improvement survey (PAIS). What patients think. *Aust Fam Physician* 2001;30:1096-2000.
- Goldstein E, Cleary PD, Langwell KM, Zaslavsky AM, Heller A. Medicare managed care CAHPS: a tool for performance improvement. *Health Care Financ Rev* 2001;22:101-7.
- Wensing M, Baker R, Szecsenyi J, Grol R; EUROPEP Group. Impact of national health care systems on patient evaluations of general practice in Europe. *Health Policy* 2004;68:353-7
- Safran DG, Karp M, Coltin K, Chang H, Li A, Ogren J, et al. Measuring patients' experiences with individual primary care physicians: Results of a statewide demonstration project. *J Gen Intern Med* 2006;21:13-21.
- Wensing M, Hermsen J, Grol R, Szecsenyi J. Patient evaluations of accessibility and co-ordination in general practice in Europe. *Health Expect* 2008;11:384-90.
- Mead N, Bower P, Roland M. General Practice Assessment Questionnaire (GPAQ)—development and psychometric characteristics. *BMC Fam Pract* 2008 doi:10.1186/1471-2296-9-13.
- Giordano LA, Elliott MN, Goldstein E, Lehman WG, Spencer PA. Development, implementation and public reporting of the HCAH PS Survey. *Med Care Res Rev* 2009 doi:10.1177/1077558709341065.
- Rosenthal M, Fernandopulle R, Ryu Song H, Bruce Landon B. Paying for quality: providers' incentives for quality improvement. *Health Aff* 2004;23:127-41.
- Robinson JC, Shortell SM, Rittenhouse DR, Fernandes-Taylor S, Gillies RR, Casalino LP. Quality-based payment for medical groups and individual physicians. *Inquiry* 2009;46:172-81.
- British Medical Association, NHS Employers. *Quality and outcomes framework. guidance for GMS contract 2009/10* (3rd rev). www.bma.org.uk/employmentandcontracts/independent_contractors/quality_outcomes_framework/qof0309.jsp (accessed 28 Jun 2009).
- Ipsos MORI. *The GP Patient Survey 2009. Technical report*. Ipsos MORI. www.gp-patient.co.uk/results/England%20Technical%20Report%2008-09.pdf (accessed 5 Jul 2009).
- Campbell J, Smith P, Nissen S, Bower P, Roland M. *Development of the national GP Patient Survey for use in primary care in the National Health Service in the UK*. BMC Health Services 2009. www.biomedcentral.com/1471-2296/10/57.
- Hays RD, Shaul JA, Williams VS, Lubalin JS, Harris-Kojetin LD, Sweeney SF, et al. Psychometric properties of the CAHPS 1.0 survey measures. *Consumer Assessment of Health Plans Study Medical Care* 1999;37(suppl 3):MS22-31.
- Groves R, Fowler F, Couper M, Lepkowski J, Singer E, Tourangeau R. *Survey methodology*. Wiley Series in Survey Methodology. 2004.
- Richards S, Coast J, Peters T. Patient reported use of health service resources compared with information from health providers. *Health Soc Care Community* 2003;11:510-8.
- Bower P, Roland M. Bias in patient assessments of general practice: General Practice Assessment Survey scores in surgery and postal respondents. *Br J Gen Pract* 2003;53:126-8.
- Mead N, Roland M. Understanding why some ethnic minority patients evaluate medical care more negatively: cross-sectional analysis of a routine patient survey in English primary care. *BMJ* 2009;339:b3450.
- Chapple A, Campbell S, Rogers A, Roland M. Users' understanding of medical knowledge in general practice. *Soc Sci Med* 2001;54:1215-24.
- Williams B, Coyle J, Healy D. The meaning of patient satisfaction: an explanation of high reported levels. *Soc Sci Med* 1998;47:1351-9.
- Gelb Safran D, Montgomery J, Chang H, Murphy J, Rogers W. Switching doctors: predictors of voluntary disenrollment for a primary physician's practice. *J Fam Pract* 2001;50:130-6.
- Rodriguez H, von Glahn T, Chang H, Rogers W, Safran DG. Measuring patients' experiences with their individual specialist physicians and their practices. *Am J Med Qual* 2009;24:35-44.

- 29 Ipsos MORI. *National GP patient surveys on access and choice summary report 2006/2007*. NHS Information Centre. www.dh.gov.uk/en/Publicationsandstatistics/PublishedSurvey/GPpatientsurvey2007/DH_075127 (accessed 28 Jun 2009).
- 30 Ipsos MORI. *GP patient survey (access) summary report 2007/2008 (England)*. NHS Information Centre. www.ic.nhs.uk/statistics-and-data-collections/supporting-information/primary-care/general-practice/gpps-2007/08/gpps-2007-08-reports (accessed 28 Jun 2009).
- 31 Checkland K, Harrison S, McDonald R, Grant S, Campbell S, Guthrie B. Biomedicine, holism and general medical practice: responses to the 2004 General Practitioner Contract. *Social Health Illness* 2008;30:788-803.
- 32 Campbell S, McDonald R, Lester H. The experience of pay for performance in English family practice: a qualitative study. *Ann Fam Med* 2008;6:228-34.
- 33 Rodríguez H, Scoggins J, von Glahn T, Zaslavsky A, Safran DG. Attributing sources of variation in patients' experiences of ambulatory care. *Med Care* 2009;47:835-41.
- 34 Groves R. Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Q* 2006;70:646-75.
- 35 Groves R, Peytcheva E. The impact of nonresponse rates on nonresponse bias: a meta-analysis. *Public Opinion Q* 2008;72:167-89.
- 36 Elliott M, Zaslavsky A, Goldstein E, Lehrman W, Hambarsoomians K, Beckett M, et al. Effects of survey mode, patient mix and non response on CAHPS hospital survey scores. *Health Serv Res* 2009;44:501-18.
- 37 Elliott MN, Edwards C, Angeles J, Hays RD. Patterns of unit and item nonresponse in the CAHPS hospital survey. *Health Serv Res* 2005;40:2096-119.
- 38 Kahn KL, Honhiu L, Adams JL, Chen WP, Tisnado DN, Carlisle DM, et al. Methodological challenges associated with patient responses to follow-up longitudinal surveys regarding quality of care. *Health Serv Res* 2003;38:1579-98.
- 39 Angus VC, Entwistle VA, Emslie MJ, Walker KA, Andrew JE. The requirement for prior consent to participate on survey response rates: a population-based survey in Grampian. *BMC Health Serv Res* 2003;3:21.
- 40 Giordano LA, Elliott MN, Goldstein E, Lehrman WG, Spencer PA. Development, implementation and public reporting of the HCAHPS survey. *Med Care Res Rev* 2009 (available online).

Accepted: 16 September 2009