

Should measures of patient experience in primary care be adjusted for case mix? Evidence from the English General Practice Patient Survey

Charlotte Paddison,¹ Marc Elliott,² Richard Parker,¹ Laura Staetsky,³ Georgios Lyratzopoulos,¹ John L Campbell,⁴ Martin Roland¹

► Additional appendices are published online only. To view these files please visit the journal online (<http://qualitysafety.bmj.com/content/21/8.toc>).

¹Cambridge Centre for Health Services Research, University of Cambridge, Cambridge, UK

²RAND Corporation, Santa Monica, California, USA

³RAND Europe, Westbrook Centre, Cambridge, UK

⁴Department of General Practice, Peninsula Medical School, Exeter, UK

Correspondence to

Professor Martin Roland, Cambridge Centre for Health Services Research, Institute of Public Health, University of Cambridge, Forvie Site, Robinson Way, Cambridge CB2 0SR, UK; mr108@cam.ac.uk

Accepted 13 April 2012

Published Online First
23 May 2012



This paper is freely available online under the BMJ Journals unlocked scheme, see <http://qualitysafety.bmj.com/site/about/unlocked.xhtml>

ABSTRACT

Objectives: Uncertainties exist about when and how best to adjust performance measures for case mix. Our aims are to quantify the impact of case-mix adjustment on practice-level scores in a national survey of patient experience, to identify why and when it may be useful to adjust for case mix, and to discuss unresolved policy issues regarding the use of case-mix adjustment in performance measurement in health care.

Design/setting: Secondary analysis of the 2009 English General Practice Patient Survey. Responses from 2163 456 patients registered with 8267 primary care practices. Linear mixed effects models were used with practice included as a random effect and five case-mix variables (gender, age, race/ethnicity, deprivation, and self-reported health) as fixed effects.

Main outcome measures: Primary outcome was the impact of case-mix adjustment on practice-level means (adjusted minus unadjusted) and changes in practice percentile ranks for questions measuring patient experience in three domains of primary care: access; interpersonal care; anticipatory care planning, and overall satisfaction with primary care services.

Results: Depending on the survey measure selected, case-mix adjustment changed the rank of between 0.4% and 29.8% of practices by more than 10 percentile points. Adjusting for case-mix resulted in large increases in score for a small number of practices and small decreases in score for a larger number of practices. Practices with younger patients, more ethnic minority patients and patients living in more socio-economically deprived areas were more likely to gain from case-mix adjustment. Age and race/ethnicity were the most influential adjusters.

Conclusions: While its effect is modest for most practices, case-mix adjustment corrects significant underestimation of scores for a small proportion of practices serving vulnerable patients and may reduce the risk that providers would 'cream-skim' by not enrolling patients from vulnerable socio-demographic groups.

BACKGROUND

Performance measurement is central to improving the quality of medical care, facilitating patient choice, and incentivising providers to improve care. Case-mix adjustment seeks to control for differences in patient characteristics not under the control of providers to facilitate fair comparison among healthcare providers by estimating the scores providers would have received if serving a common population.¹ Such adjustment can be applied to measures of clinical process,^{2–3} patient experience,^{1 4–7} and medical outcomes.^{8–11}

The use of case-mix adjustment for mortality is relatively well established^{12–15}—although uncertainty exists about optimal methods^{10 16}—while the application of case-mix adjustment for clinical process and patient experience measures varies within the UK and internationally. For example, in the USA the same Consumer Assessment of Healthcare Providers and Systems (CAHPS) data are presented by the Centers for Medicare and Medicaid Services (CMS) adjusted,¹⁷ and by the National Center for Quality Improvement (NCQA) unadjusted.¹⁸

Variability in the use of case-mix adjustment may reflect both ideological differences and practical uncertainties about when and how to adjust for patient case mix. Some have suggested that performance indicators should not be case-mix adjusted, citing concerns that such an approach disregards the imperative to provide the best possible care to all patients. Yet organisations with high proportions of patients who are difficult to treat or generally more likely to report negative experiences could be disadvantaged by performance

incentives based on unadjusted performance measures. Such disadvantage may encompass loss of income (eg, under pay-for-performance), but may also damage staff morale, and undermine the ability to recruit and retain high-calibre healthcare professionals and/or attract and retain patients.

Further empirical research is needed to assist and inform such decision making, given the inconsistent use and ongoing debate about the benefits of case-mix adjustment.^{5 19 20} In England, detailed measurement of clinical performance is used as part of a pay-for-performance scheme in primary care,²¹ and between 2008 and 2011, the scheme included financial incentives based on data from a national survey of patient experience.^{22 23} Using data from this survey, we quantify the impact of case-mix adjustment on the mean scores and rankings of primary care practices. We ascertain who would be the winners and losers under case-mix adjustment, by how much, and why. Our aim is to demonstrate the difference which case-mix adjustment might make in performance assessment, to identify why and when it may be useful, and to discuss important policy issues in case-mix adjustment.

METHODS

Anonymised data were analysed from all 2 163 456 people registered with 8267 primary care practices in England who responded to the 2009 General Practice Patient Survey, a national survey of patient experience with primary care (response rate 38%). Women, middle-aged patients, and those in affluent areas were more likely to respond to the survey, consistent with response patterns reported for other patient experience surveys.²⁴ Details of the survey and method of administration have been published elsewhere.^{23 25}

Patient experience measures

We measured patient experience using a single item to assess overall satisfaction with primary care services, and 12 items assessing patient experience in three domains of primary care: access, interpersonal care, and anticipatory care planning. Response options included both dichotomous categories (yes/no) and three-point, four-point and five-point Likert scales. The questionnaire can be viewed at <http://www.gp-patient.co.uk/questionnaires/> and the content of each question is summarised in the appendix, table A1. All measures were linearly re-scaled to a 0–100 range (most favourable response =100) to facilitate comparisons across questions. Composite measures were computed separately for doctor communication and nurse communication using the mean of non-missing items from all respondents answering four or more of the seven items which made up each composite.

Case mix

Five patient characteristics were investigated as potential case-mix adjusters: age (eight ordinal categories from 18–24 to 85+; 55–64 (reference group)), male gender, Office of National Statistics ethnicity (white (reference group); mixed; South Asian; black; Chinese; other), self-reported health (excellent (reference group), very good, good, fair, poor), and socio-economic status of residential address (by quintile; using the highest level of deprivation as the reference group). The last of these was the only variable measured at aggregate level and was derived by linking patient postal codes to the 2007 Lower Super Output Area Index of Multiple Deprivation²⁶ (higher quintiles represent greater socio-economic deprivation).

Analyses

We used linear mixed effects regression models with practice as a random effect and case-mix adjusters as fixed effects, to predict each of 13 measures of patient experience with primary care (multiple items for access, interpersonal care, anticipatory care planning, and a single item measuring overall satisfaction with primary care services). The use of practice random effects ensures that adjustments are limited to differences observed within practices and are not biased by the confounding of overall practice performance with between-practice differences in patient characteristics.

Metrics

Overall impact of case-mix adjustment

To quantify the impact of case-mix adjustment we used two primary metrics:

(1) To convey the absolute change resulting from case-mix adjustment we calculated differences on the 0–100 scale between case-mix adjusted and unadjusted practice-level mean scores. We described absolute changes in practice means resulting from adjustment by showing the distribution of the changes in means, and by calculating the proportion of variance in adjusted practice-level means attributable to case-mix adjustment, $p=(1-r^2)$ where r is the Pearson's correlation coefficient between adjusted and unadjusted practice-level means.

(2) To convey relative change resulting from case-mix adjustment we calculated changes in practice rankings. We described the frequency of large impacts by counting the proportion of practices that change rank by 10 or more and by 20 or more percentile points after case-mix adjustment was applied. We also used Kendall's Tau (τ) to calculate the discordance of ranks under case-mix adjustment, d , where $d=(1-\tau)/2$, represents the probability that the higher rank of a given practice relative to another practice becomes the lower rank relative to the same practice after case-mix adjustment (or vice versa).

How the effect of case-mix adjustment varies in relation to practice characteristics

To identify the type of practices likely to gain or lose most as a result of case-mix adjustment, we first arranged all practices in descending order according to the impact of case-mix adjustment on practice-level means (adjusted–unadjusted) for each of the two questions, 7 and 10, which related to the ability of patients to get appointments with their general practitioner and which were associated with pay-for-performance between 2008 and 2011. We then calculated descriptive characteristics for respondents in practices in the highest decile (the biggest ‘winners’ from case mix adjustment); the lowest decile (the biggest ‘losers’); and middle 10% of practices (‘typical impacts’).

Relative importance of individual case-mix adjustors

To measure the contribution of individual case-mix adjustment variables we used discordance in ranks to describe the *total* contribution of a single case-mix adjustor if it were the only adjustor employed (assessed by comparing ranks from models adjusting only for the case-mix adjustor being examined to unadjusted ranks), and the *unique* contribution of a single case-mix adjustor above and beyond all others (assessed by comparing ranks from fully adjusted models to ranks from models adjusting for all but the one adjustor being examined).

The impact of any case-mix adjustor is a function of two factors: how strongly predictive the adjustor is of individual patient scores within practices; and how much the adjustor varies between practices.⁴ To assess the predictive strength of each adjustor we examined standardised coefficients showing the relationship between individual case-mix variables and each of the 13 measures of patient experience in our primary regression models. We then calculated practice-level intraclass correlation coefficients (ICCs) for each adjustor to measure its heterogeneity between practices.

SAS V.9.2 was used to perform all linear mixed effects modelling. SPSS/PASW for Windows (V.18.0.3) and R V.2.12.0 were used for all other analyses.

RESULTS

Demographic characteristics and the self-reported health of the 2 163 456 respondents to the 2009 General Practice Patient Survey appear in **table 1**. Fifty-eight per cent of respondents were women; 30% were aged over 65; and 14% were non-white. Practice-level and patient-level mean scores for questions from the General Practice Patient Survey appear in appendix table A1.

Overall impact of case-mix adjustment

Table 2 shows the impact of case-mix adjustment on practice-level mean scores and practice ranks. Case-mix

Table 1 Demographic characteristics and self-reported health of respondents to the 2009 General Practice Patient Survey (England)

	Total survey respondents (n)	% of survey respondents
Gender		
Male	890 241	42.4
Female	1 207 171	57.6
Age group		
18–24	103 040	4.9
25–34	229 546	10.9
35–44	332 017	15.8
45–54	374 722	17.8
55–64	426 786	20.3
65–74	349 759	16.6
75–84	220 795	10.5
85+	64 943	3.1
Ethnic group		
White	1 809 150	86.3
Mixed	15 080	0.7
Asian*	112 905	5.4
Black†	57 754	2.8
Chinese	9759	0.5
Other ethnic group	90 644	4.3
Socio-economic deprivation quintile		
1 (affluent)	378 798	17.5
2	410 445	19.0
3	423 179	19.6
4	440 427	20.4
5 (deprived)	506 286	23.4
Self-rated health		
Excellent	194 735	9.5
Very good	610 217	29.6
Good	737 993	35.8
Fair	398 319	19.3
Poor	118 102	5.7

*Indian, Pakistani, Bangladeshi, any other Asian.

†Black Caribbean, Black African, any other Black.

adjustment accounted for 1–12% of the variance in practice-level mean scores (see column 1), and changed the relative ranks of between 2% and 11% of practices in pairwise comparisons (see column 2). The impact of case-mix adjustment on practice-level means varied by patient experience domain, with a larger impact on anticipatory care planning items (questions 28 and 29) and a smaller impact for questions relating to access (getting through on the telephone; getting an urgent appointment; booking ahead; and evaluation of waiting time).

Adjusting for patient characteristics had modest impacts for most practices, but between 0.4% and 29.8% of practices changed rank by 10 percentile points or more (**table 2**, columns 3 and 4), and up to 6% of practices changed rank by 20 percentile points or more (**table 2**, columns 5 and 6). As can be seen in **table 2**, gains of 10 or more percentile points of rank were somewhat more common than losses of 10 or more percentile points.

Table 2 Impact of adjustment for all case-mix variables on practice level mean scores, and practice rankings

General practice patient survey question	Percentage of practice score variance due to case-mix adjustment† (column 1)	Percentage reranking among practices‡ (column 2)	Practices losing 10+ percentile ranks (%) (column 3)	Practices gaining 10+ percentile ranks (%) (column 4)	Practices losing 20+ percentile ranks (%) (column 5)	Practices gaining 20+ percentile ranks (%) (column 6)
Helpfulness of receptionists						
Q4*	5.0	6.4	3.3	6.1	0.1	1.6
Getting through on the phone						
Q5a	0.7	2.4	0.0	0.4	0.0	0.0
Ability to get urgent appointment						
Q7	2.0	4.8	0.7	3.4	0.0	0.3
Ability to book ahead						
Q10	1.7	4.2	0.1	2.4	0.0	0.1
Evaluation of surgery waiting time						
Q14	4.8	6.6	2.8	6.8	0.0	1.8
Seeing the doctor you prefer						
Q16	6.3	7.4	3.9	9.2	0.0	2.5
Satisfaction with opening hours						
Q17	5.9	7.7	7.6	8.0	0.4	1.5
Doctor patient communication						
Q20	5.3	7.0	4.5	7.1	0.1	2.0
Confidence and trust in doctor						
Q21	5.3	7.3	5.4	7.6	0.2	1.8
Nurse patient communication						
Q24	7.2	7.7	5.8	8.2	0.1	2.8
Overall satisfaction						
Q25	6.2	7.2	4.6	7.5	0.2	2.0
Did doctor or nurse agree a care plan?						
Q28	12.3	11.1	15.1	14.7	1.8	4.9
Has care plan helped improve care?						
Q29	10.8	10.1	12.3	12.5	1.5	3.7

*General Practice Patient Survey question number.

†Percentage of variation in practice-level means due to case-mix adjustment calculated as $1 - r^2$ where r is the Pearson's correlation coefficient between adjusted and unadjusted practice-level means.

‡Calculated $((1 - \tau)/2) \times 100$ where τ is Kendall's tau correlation coefficient between the adjusted practice level means and unadjusted practice level means with shrinkage.

Figure 1 shows the magnitude and direction of differences in practice-level mean scores on 13 patient experience questions after case-mix adjustment is applied. The proportion of practices that are outliers (more than 1.5 times the IQR outside the actual IQR) ranges from 2.5% to 6.9% across questions. As can be seen in figure 1 these differences in practice-level means are generally 5–10 points on the 0–100 scale (corresponding to 20–50 percentile points of change in rank, see appendix, figure A1), with the larger outlier adjustments almost entirely positive, that is, representing practices whose mean scores increased as a result of case-mix adjustment.

How the effect of case-mix adjustment varies in relation to practice characteristics

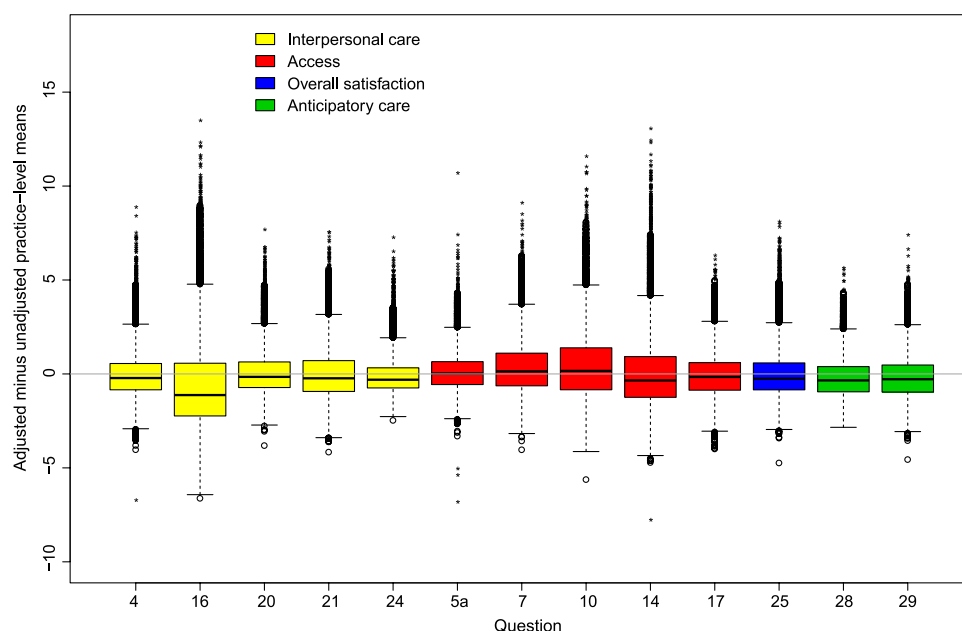
Practices showing the largest gain in scores from case-mix adjustment ('winners' in the top decile), middle decile practices ('typical', very small adjustments), and

bottom decile practices ('losers' from adjustment), had very different average patient populations, especially with respect to deprivation and ethnicity (see table 3). Among the 10% of practices gaining most from case-mix adjustment, on average, 82% of patients lived in the most deprived areas (lowest two quintiles), 45% of patients were white and 31% South Asian. In comparison, the 10% of practices who lost most from case-mix adjustment had a patient population that averaged 97% white, and only 10% of their patients lived in the most deprived areas. Practices gaining most from case-mix adjustment were on average smaller and treated patients who tended to be younger and in worse health compared with other practices.

Relative importance of individual adjustors

Figure 2 summarises the relative importance of each of the five case-mix adjustor variables, averaged across all items. Age is the most influential adjustor (accounting

Figure 1 Boxplot showing impact of case-mix adjustment on practice level means (adjusted–unadjusted means). Whiskers show 1.5 IQR beyond the actual IQR; cases outside the IQR by at least 3 IQR are shown as*.



for about 4% discordance by itself) and gender the least (~1% discordance). Ethnicity is nearly as important as age when taken alone, but its unique contribution is similar to those for deprivation and self-reported health (~2%). The ICCs show that deprivation and ethnicity vary greatly across practices and gender almost not at all. Standardised regression coefficients (appendix table A2) show relatively strong relationships of age and self-reported health with measures of patient experience within practices, but relatively small coefficients for gender and deprivation.

DISCUSSION

In a study of 2 163 456 respondents from 8267 English primary care practices we show that case-mix adjustment of practice-level scores results in relatively few large

adjustments (which were mainly positive), and many small adjustments (which were more often negative). However, the largest effects were on a distinct subset of practices whose patients were more likely than average to be from South Asian or other ethnic minorities, young, in poorer health, and living in deprived areas.

Although only a small number of practices would benefit significantly from case-mix adjustment, we propose that such adjustment should be applied because it meaningfully improves performance measurement for practices with less typical and often under-privileged patient populations. This would discourage practices from ‘cream-skimming’ by avoiding enrolling patients who could be seen as ‘hard to treat’, and increase perceptions of fairness and engagement in quality improvement activities.

In a study of 27 practices in the UK (patient race/ethnicity 97% white), Salisbury and colleagues conclude

Table 3 Characteristics of practices in the top 10% (gain most from case-mix adjustment), mid 10%, and bottom 10% when all practices are arranged by magnitude of change in practice-level mean scores after case-mix adjustment

Description of patient and practice characteristics	Top 10% of practices*	Mid 10% of practices	Bottom 10% of practices
Male (%)†	46	42	42
Aged <35 (%)	30	15	9
White (%)	45	93	97
Average percentage fair or poor self-reported health	30	25	20
Average percentage IMD quintile 4 or 5 (deprived)	82	43	11
Average practice size (registered population)	3426	5072	5434
Impact of case-mix adjustment on practice-level means for two questions used in pay-for-performance scheme			
Ability to get urgent appointment‡	3.3	0.2	–1.4
Ability to book ahead‡	4.3	0.2	–2.1

*When all practices arranged in deciles by order of magnitude of change in practice-level mean score after case-mix adjustment for Q7 and Q10, with top 10% showing greater benefit from adjustment.

†Average for practices in named decile block (top, mid, or bottom 10%).

‡Adjusted minus unadjusted practice-level mean.

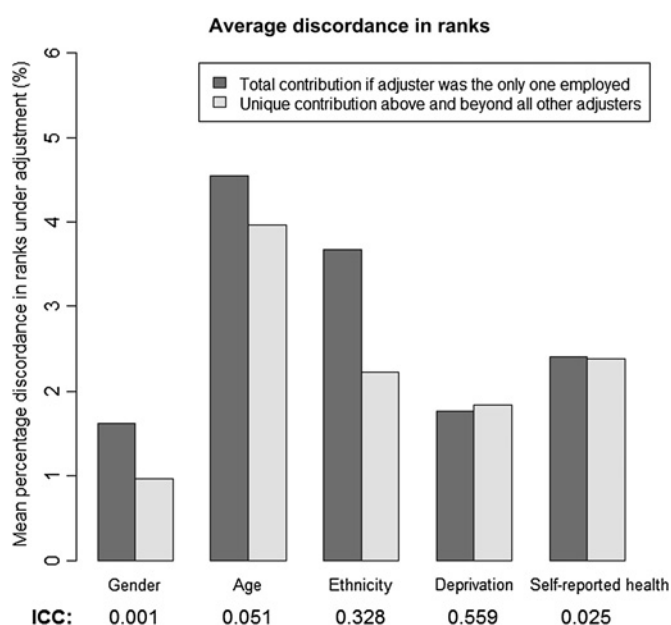


Figure 2 Average discordance of ranks across all items using data on Kendall's Tau and intraclass correlation coefficients (ICCs) to identify the variables that have the biggest impact when case-mix adjustment is applied.

that 'adjusting for patients' characteristics makes very little difference to practices' scores or to the performance of individual practices relative to other practices'.⁶ Our study, with a larger sample of practices and including a broader mix of patient characteristics (practice N=8267; patient race/ethnicity 14% non-white), suggests that case-mix adjustment has a non-trivial impact on the assessment of performance for practices serving less typical, and often disadvantaged, patient populations. Unfairly disadvantaging such practices in performance measurement could have negative implications for retention and recruitment of healthcare staff working in them, and for attracting and retaining patients.

Our study builds on what is known from previous research by addressing some of the limitations of previous studies, which include smaller sample sizes, non-random selection of practices, and a low proportion of non-white respondents.⁶ A particular strength of our study is the large sample size, enabling us to investigate the impact of case-mix adjustment on practices serving less typical patient populations. Because of the population basis of the sampling, our study is also not subject to problems such as regional variation in coding which may cause problems when using diagnostic coding for case-mix adjustment.²⁷

Although parsimonious models are often desirable, one limitation of this study is the relatively small number (five) of case-mix adjusters included. While age, gender, ethnicity, self-reported health, and socio-economic deprivation are considered pertinent for case-mix adjustment in

both the UK and the USA there are additional patient characteristics, such as language spoken at home, that have been used as case-mix adjusters in prior research⁵ but were not measured in this study. Another limitation is the modest response rate to the survey (38%). However in our previous analysis of two questions associated with payment to practices we found minimal evidence of non-response bias,²³ and this is consistent with a meta-analysis of survey methodology literature showing that response rates are only weakly associated with non-response bias among studies employing methodology similar to ours.²⁸ There are limits to the generalisability of our findings. The impact of individual case-mix adjusters may vary between countries due to differences in, for example, the constituent racial/ethnic groups in each nation, and we were not able in our study to explore variation across nations or healthcare settings.

Implications for health policy and practice

Because case-mix adjustment reduces bias^{4 5} and improves the validity of performance measurement (especially for some practices) it is integral to supporting patient choice and facilitating quality improvement in hospitals and in primary care.⁵ In addition, by improving *perceived* fairness, or face validity,²⁹ case-mix adjustment helps to maintain the credibility of pay-for-performance schemes, and focuses conversations on how to improve patient experiences⁴ by avoiding arguments from providers that 'my population is special'.²⁹ However, care must still be taken to communicate that negative adjustments are not 'penalties', and we recognise, like others,^{2 30} that there is a risk that case-mix adjustment could remove incentives for providers to address disparities in care and institutionalise substandard care by 'masking' poor care provided to some patient subgroups. In order to make visible any disparities in the provision of care by socio-economic status or race/ethnicity and to minimise the risk of institutionalising substandard care, it is important that case-mix adjustment is used in conjunction with strategies that collect data to monitor healthcare disparities and report quality measures stratified by, for example, socio-economic position and race/ethnicity.³⁰

Deciding when and how to adjust performance scores for case mix

Case-mix adjustment is most useful when patient characteristics vary substantially between providers, and where these patient characteristics are strongly related to performance measures.⁵ Combined with previous US findings,^{5 7} our data suggest that age and health status are broadly important adjusters across nations and healthcare settings. Our research also suggests that adjustment for race/ethnicity may be important to ensure equitable comparison, and, when results are linked to financial

incentives, equitable distribution of resources. Socio-economic status/deprivation was less important in these data than is often the case in the USA; these differences may reflect differences in healthcare systems or in the measurement of socio-economic status.

CONCLUSIONS

Case-mix adjustment has been inconsistently applied to performance measures in healthcare both in the UK and internationally. Our results show that the impact of case-mix adjustment on practice-level performance scores is modest for most practices but may meaningfully improve the measurement of performance for practices with less typical patient populations, discouraging practices from 'cream-skimming' by avoiding enrolling patients who could be seen as 'hard to treat'. Without adjustment for patient mix, reports of organisational performance may disadvantage those practices serving less typical, often disadvantaged, patient populations.

Contributors All authors have made substantial contributions to conception and design, and analysis and interpretation of data; and drafting and revising the article. All authors have approved the final version. CP is the guarantor of the paper.

Funding The study was funded by a grant from the UK Department of Health. The views expressed are those of the authors and not of the Department of Health.

Competing interests The authors have no conflicts of interest as defined by the International Committee of Medical Journal Editors. MR and JC act as academic advisors to Ipsos MORI for the development of the General Practice Patient Survey and have received remuneration for this.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement Individual level data were provided to us by Ipsos MORI with a covering confidentiality agreement with the Department of Health. Under this agreement we are not at liberty to share the dataset with third parties.

REFERENCES

- Elliott MN, Zaslavsky AM, Goldstein E, *et al.* Effects of survey mode, patient mix, and nonresponse on CAHPS® hospital survey scores. *Health Serv Res* 2009;44:501–18.
- Hong CS, Atlas SJ, Chang Y, *et al.* Relationship between patient panel characteristics and primary care physician clinical performance rankings. *JAMA* 2010;304:1107–13.
- Mehta RH, Liang L, Karve AM, *et al.* Association of patient case-mix adjustment, hospital process performance rankings, and eligibility for financial incentives. *JAMA* 2008;300:1897–903.
- Hargraves JL, Wilson IB, Zaslavsky A, *et al.* Adjusting for patient characteristics when analyzing reports from patients about hospital care. *Med Care* 2001;39:635–41.
- O'Malley AJ, Zaslavsky AM, Elliott MN, *et al.* Case-mix adjustment of the CAHPS hospital survey. *Health Serv Res* 2005;40:2163–81.
- Salisbury C, Wallace M, Montgomery AA. Patients' experience and satisfaction in primary care: secondary analysis using multilevel modelling. *BMJ* 2010;341:c5004.
- Zaslavsky AM, Zaboriski L, Ding L, *et al.* Adjusting performance measures to ensure equitable plan comparisons. *Health Care Financ Rev* 2001;22:109–26.
- Bottle A, Aylin P. Intelligent information: a national system for monitoring clinical performance. *Health Serv Res* 2008;43:10–31.
- Lingsma HF, Steyerberg EW, Eijkemans MJ, *et al.* Comparing and ranking hospitals based on outcome: results from The Netherlands Stroke Survey. *QJM* 2009;103:99–108.
- Mohammed MA, Deeks JJ, Gilling A, *et al.* Evidence of methodological bias in hospital standardised mortality ratios: retrospective database study of English hospitals. *BMJ* 2009;338:b780.
- Smith AW, Reeve BB, Bellizzi KM, *et al.* Cancer, comorbidities, and health-related quality of life of older adults. *Health Care Financ Rev* 2008;29:41–56.
- Jarman B, Pieter D, van der Veen AA, *et al.* The hospital standardised mortality ratio: a powerful tool for Dutch hospitals to assess their quality of care? *Qual Saf Health Care* 2010;19:9–13.
- Pine M, Jordan HS, Elixhauser A, *et al.* Enhancement of claims data to improve risk adjustment of hospital mortality. *JAMA* 2007;297:71–6.
- Popescu I, Vaughan-Sarrazin MS, Rosenthal GE. Differences in mortality and use of revascularization in black and white patients with acute MI admitted to hospitals with and without revascularization services. *JAMA* 2007;297:2489–95.
- Sharek PJ, Parast LM, Leong K, *et al.* Effect of a rapid response team on hospital-wide mortality and code rates outside the ICU in a Children's Hospital. *JAMA* 2007;298:2267–74.
- Shahian DM, Wolf RE, Iezzoni LI, *et al.* Variability in the measurement of hospital-wide mortality rates. *N Engl J Med* 2010;363:2530–9.
- National Quality Measures Clearinghouse. *Hospital Inpatients' Experiences: Adults Inpatients' Ratings of This Hospital*. Rockville, MD: National Quality Measures Clearinghouse. <http://www.qualitymeasures.ahrq.gov/content.aspx?id=26585> (accessed 8 Mar 2012).
- NCQA. *HEDIS. Volume 3: Specifications for Survey Measures*. Washington, DC: NCQA, 2012. <http://www.ncqa.org/tabid/1415/Default.aspx> (accessed 4 May 2012).
- Kontopantelis E, Roland M, Reeves D. Patient experience of access to primary care: identification of predictors in a national patient survey. *BMC Fam Pract* 2010;11:61.
- Mead N, Roland M. Understanding why some ethnic minority patients evaluate medical care more negatively than white patients: a cross sectional analysis of a routine patient survey in English general practices. *BMJ* 2009;339:b3450.
- Roland M. Linking physician pay to quality of care: a major experiment in the UK. *N Engl J Med* 2004;351:1448–54.
- Campbell J, Smith P, Nissen S, *et al.* The GP Patient Survey for use in primary care in the National Health Service in the UK—development and psychometric characteristics. *BMC Fam Pract* 2009;10:57.
- Roland M, Elliott M, Lyrtzopoulos G, *et al.* Reliability of patient responses in pay for performance schemes: analysis of national General Practitioner Patient Survey data in England. *BMJ* 2009;339: b3851.
- Elliott MN, Edwards C, Angeles J, *et al.* Patterns of unit and item non-response in the CAHPS Hospital Survey. *Health Serv Res* 2005;40:2096–119.
- Ipsos MORI. *Technical Annex for the GP Patient Survey*. London: Ipsos MORI, 2010.
- Department for Communities and Local Government. *The English Indices of Deprivation*. London: Department for Communities and Local Government, 2010. <http://www.communities.gov.uk/publications/corporate/statistics/indices2010> (accessed 8 Mar 2012).
- Welch HG, Sharp SM, Gottlieb DJ, *et al.* Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. *JAMA* 2011;305:1113–18.
- Groves RM, Peytcheva E. The impact of nonresponse rates on nonresponse bias. *Public Opin Q* 2008;72:167–89.
- Eselius LL, Cleary PD, Zaslavsky AM, *et al.* Case-mix adjustment of consumer reports about managed behavioral health care and health plans. *Health Serv Res* 2008;43:2014–32.
- Fiscella K, Franks P, Gold MR, *et al.* Inequality in quality: addressing socioeconomic, racial, and ethnic disparities in health care. *JAMA* 2000;283:2579–84.

**Should measures of patient experience in primary care be adjusted for case mix?
Evidence from the English General Practice Patient Survey.**

Paddison et al (corresponding author: Roland).

Appendix: Supplementary tables and figures

Appendix Table A1. Practice-level and patient-level means for 13 patient experience items (including two composite measures)

General Practice Patient Survey question	Practice level			Patient level	
	Mean	Standard deviation	Minimum	Maximum	Mean
Helpfulness of receptionists	83.5	6.6	0.0	100.0	83.4
Getting through on the phone	69.5	13.4	22.1	100.0	69.0
Ability to get urgent appointment	84.8	10.7	0.0	100.0	85.1
Ability to book ahead	77.6	15.4	0.0	100.0	77.8
Evaluation of surgery waiting time	82.7	10.2	30.2	100.0	82.7
Seeing the doctor you prefer	76.9	11.4	30.4	100.0	76.6
Satisfaction with opening hours	79.7	5.5	50.0	100.0	79.7
Doctor patient communication (composite measure)	83.5	5.7	46.4	100.0	83.7
Confidence and trust in doctor	84.1	6.6	49.8	100.0	84.3
Nurse patient communication (composite measure)	87.0	4.5	58.3	100.0	87.3
Overall satisfaction	86.0	5.9	0.0	100.0	86.1
Did doctor or nurse agree a care plan?	87.6	5.7	0.0	100.0	88.1
Has care plan helped improve care?	73.3	6.1	0.0	100.0	73.7

Appendix Table A2. Standardized coefficients for case-mix adjusters age, gender, ethnicity, deprivation, and self-reported health (continued overleaf)

General Practice Patient Survey question	Gender (reference group female)	Age (reference group 55-64)								Ethnicity (reference group White)			
		18-24	25-34	35-44	45-54	65-74	75-84	85 or older	Mixed	Black	Asian	Chinese	Other
Helpfulness of receptionists	0.14	-10.37	-8.66	-5.58	-3.71	4.95	8.29	8.75	0.31	2.92	-3.75	-4.01	-0.55
Getting through on the phone	0.83	-7.41	-5.35	-4.00	-3.02	3.44	5.91	6.66	0.45	1.31	-2.89	-3.22	-0.02
Ability to get urgent appointment	-0.82	-7.57	-4.11	-2.52	-3.20	3.81	5.82	6.19	-2.08	-0.85	-3.62	-0.35	-1.26
Ability to book ahead	2.69	-7.63	-4.41	-3.66	-3.17	3.33	6.11	7.40	-0.82	-0.73	-5.57	-0.08	-1.14
Evaluation of surgery waiting time	1.61	-12.15	-10.11	-8.00	-4.26	4.02	6.39	6.24	-1.63	-0.48	-7.17	-6.57	-2.42
Seeing the doctor you prefer	1.99	-15.20	-13.50	-10.81	-6.11	5.09	7.49	6.91	-3.05	-4.93	-5.46	-8.62	-3.96
Satisfaction with opening hours	-1.45	-8.74	-8.06	-5.83	-4.77	5.95	7.42	5.93	0.11	2.67	-2.55	-6.40	-0.38
Doctor patient communication	-0.23	-8.94	-8.01	-4.79	-2.70	2.89	3.87	3.29	-1.58	0.04	-2.79	-6.59	-2.74
Confidence and trust in doctor	0.74	-10.45	-9.61	-4.98	-2.73	3.49	5.17	4.91	-1.99	-0.43	-1.85	-8.41	-3.86
Nurse patient communication	-0.20	-5.09	-3.95	-2.35	-1.60	1.96	2.95	2.81	-1.92	-1.29	-5.07	-8.04	-3.10
Overall satisfaction	-0.22	-9.92	-8.87	-5.46	-3.35	3.87	5.39	4.52	-1.09	1.16	-4.19	-7.36	-3.09
Did doctor or nurse agree a care plan?	0.56	-12.99	-11.10	-6.07	-3.27	2.88	3.18	2.40	-1.67	0.58	-3.39	-1.81	-2.29
Has care plan helped improve care?	1.17	-14.95	-12.61	-7.63	-4.26	4.53	5.03	3.50	-0.84	0.03	-4.46	-0.66	-1.44

Appendix Table A2. (continued) Standardized coefficients for case-mix adjusters age, gender, ethnicity, deprivation, and self-reported health

General Practice Patient Survey question	Deprivation (reference group 5, most deprived)				Self-reported health (reference group excellent)			Practice random effects variance	Residual variance
	1	2	3	4	Very good	Good	Fair		
Helpfulness of receptionists	-2.10	-1.76	-1.50	-1.09	-3.06	-5.00	-5.92	34.83	405.58
Getting through on the phone	-2.48	-2.08	-1.70	-1.21	-4.46	-7.75	-9.74	172.45	642.49
Ability to get urgent appointment	0.62	0.43	0.15	-0.18	-2.28	-4.56	-6.11	94.29	1,150.86
Ability to book ahead	1.32	0.77	0.71	0.17	-2.67	-5.45	-7.52	213.65	1,488.30
Evaluation of surgery waiting time	-0.20	-0.09	-0.07	-0.15	-2.48	-4.98	-7.40	82.27	759.01
Seeing the doctor you prefer	0.16	0.48	0.35	0.23	-2.95	-4.76	-5.15	111.09	753.32
Satisfaction with opening hours	-3.65	-2.98	-2.53	-1.63	-3.80	-6.43	-7.11	24.48	496.96
Doctor patient communication	-0.65	-0.53	-0.52	-0.36	-3.75	-7.11	-8.69	23.56	322.17
Confidence and trust in doctor	-0.50	-0.54	-0.64	-0.49	-4.02	-7.86	-10.20	31.83	649.91
Nurse patient communication	-0.97	-0.82	-0.70	-0.48	-3.39	-6.44	-7.57	12.88	243.52
Overall satisfaction	-1.42	-1.18	-1.05	-0.77	-3.36	-6.56	-8.52	23.82	328.04
Did doctor or nurse agree a care plan?	-0.32	-0.21	-0.24	-0.50	-2.61	-5.03	-8.57	9.81	1,014.85
Has care plan helped improve care?	-2.44	-1.79	-1.50	-1.05	-7.09	-11.47	-15.43	15.47	949.56

Appendix Figure A1. Boxplot showing impact of case-mix adjustment on change in percentile ranks
Whiskers show 1.5 Inter-quartile ranges (IQR) beyond the actual IQR;
Cases outside the IQR by at least 3 IQR are shown as *.

