ANNALS OF FAMILY MEDICINE

Online Supplementary Material

Llanwarne NR, Abel GA, Elliott MN, et al. Relationship between clinical quality and patient experience: analysis of data from the English Quality and Outcomes Framework and the national GP Patient Survey. *Ann Fam Med*. 2013;11(5):467-472.

http://www.annfammed.org/content/11/5/467

Supplemental Appendix. Statistical Methods

The General Practice Patient Survey (GPPS)[1] is sent to 5.5 million patients in England each year. It is sent as a postal questionnaire to patients who have been continuously registered with a family practice for at least 6 months. It can also be answered online or on the telephone in 15 languages. Stratified random samples are drawn from general practice lists resulting in an average of 260 patients respondents per practice. The sampling procedure involves oversampling from small practices and practices with low response rates in previous surveys to obtain similar numbers of respondents from each practice. Further details of the survey and its development can be found elsewhere.[2,3] The analysis reported here uses data from year 4 of the survey (April 2009 to March 2010), where 2,169,718 patients responded (response rate 39%).

The GPPS has 29 questions addressing a range of issues including access to care (eg, ease of getting appointments); doctor, nurse, and receptionist communication; trust in doctors; ability to get continuity of care; experience of care planning; and overall satisfaction. The questionnaire also includes questions about the patient including age, sex, ethnicity, and state of health. Additionally, information is included on the Index of Multiple Deprivation score of the Lower Super Output Area of patients' residence.[4] The questions from the GPPS analyzed in this study are listed in Table 2 in the article.

The Quality and Outcomes Framework (QOF) is the annual incentive program for family practices in England and has been in place since 2004.[5] Currently QOF awards practices achievement points in 4 domains (clinical, organizational, patient experience, and additional services). In 2009/2010 the clinical domain contained 89 indicators concerning 19 different conditions. Although some of the clinical indicators relate to structural aspects of service quality (eg, the presence or absence of specific disease registries), most (71/89) relate to processes of care (eg, the percentage of patients with a condition who receive the appropriate treatment or achieve set levels of control for risk factors, such as blood pressure or cholesterol). Practices can exclude individual patients from the data used to calculate QOF scores for whom treatment is judged to be contraindicated or inappropriate.[6] QOF data are extracted automatically from general practice electronic medical records and practice level QOF performance, prevalence data and exception reporting are all published on the NHS Information Centre (IC) website.[7]

For each practice patient numbers, broken down by sex and age-group, were provided by the NHS Information Centre for 2009.

Data Preparation

The GPPS questions were first linearly rescaled on a 0 to 100 scale. In the case of the doctor and nurse communication items, a composite was calculated as the mean of 7 subitems when at least 4 had been answered. We then obtained practice-level scores by calculating shrunken case-mix adjusted estimates from linear regression with a random practice effect. Case-mix adjustment variables were age, sex, ethnicity, deprivation, and self rated health, as included in the GPPS dataset. Shrunken, or empirical Bayes, estimates of practice level means are generally considered to provide the best unbiased predictions. The alternative is to use maximum likelihood estimates, but these would tend to produce more extreme estimates for practices with small numbers of GPPS responders.

For each of the clinical process indicators within QOF, we created a summary score as follows. First, we added the patients excluded via exception reporting to the denominator used for each indicator to give a measure of population achievement. Second, for each indicator we obtained a shrunken estimate of the proportion of patients for whom the measure was met in each practice from a logistic regression with no fixed-effect explanatory variables but including a random practice effect. Ideally, we would like to include case-mix variables in these models; however, the QOF data are not recorded at the patient level and so doing so is not possible. Third, a weighted mean of the shrunken proportions was calculated for each clinical factor, where the weights were equal to the number of QOF points attributed to the individual indicators. These weights were used, as they represent an independent judgment of the relative importance of each indicator. An overall clinical summary score was also calculated in the same way using all appropriate indicators within a clinical domain. Similarly we have organized the QOF indicators by type of indicator (recording/review, tests/specialist referral, treatment, intermediate outcome) and created a summary score for each type in the same way.

Our method for creating QOF practice level summary indicators has limitations. The creation of shrunken estimates for each indicator that are then combined is strictly valid only under the assumption that the score in each indicator is independent of the other indicators. Such is not the case, as a number of the same patients will have been included in the same indicators within a clinical domain (and to a lesser extent across domains). As stated above, patient-level QOF data are not available, and so it is impossible to determine the level of correlation between indicators at the patient level. The result is that the summary scores we created are shrunk by an excessive but unknown degree. While this is less than perfect, it will only tend to reduce variation and thus underestimate any association.

As a sensitivity analysis, we repeated all of the analysis outlined below using summary QOF scores calculated from observed proportions rather than shrunken proportions.

We calculated a number of items describing the demographic composition (case-mix) of each practice. Practice level proportions of male and female patient numbers were calculated from the general practice census data. Shrunken estimates of proportion of black, Asian, Chinese, mixed race, and other (nonwhite) ethnicities were obtained from the GPPS data using individual logistic regression models with a random practice effect and no fixed effects. A shrunken estimate of the mean deprivation score of the practice population was obtained from a linear mixed model with a random practice effect and no fixed effects.

Analysis

We looked at the crude association between the different QOF indicators and the GPPS scores of interest by calculating the Spearman's rank correlation coefficient between each QOF summary score and each GPPS shrunken practice score. To account for multiple testing, we applied Bonferroni corrections to the *P* values. Practices with fewer than 100 responses to the GPPS were excluded, as were practices with fewer than 30 patients with cancer, dementia, epilepsy, heart failure, mental health, cardiovascular disease-primary prevention, and 100 patients for all other conditions. These exclusions aim to remove practices where there is large uncertainty in the estimation of the practice means and proportions because of small sample sizes. Including data in our analysis containing large uncertainty will result in an underestimate of the strength of associations measured.

To investigate the relationship between QOF scores and GPPS scores in more detail, we performed a series of linear regressions with the QOF clinical summary score as the outcome variable. In each model 1 GPPS score was an explanatory variable along with either (1) no other fixed effects, or (2) practice population demographics. We also performed a number of linear regressions to calculate the proportion of variance in the QOF clinical summary score explained by different groups of GPPS questions that looked at different area of patient experience (access, continuity of care, communication, overall satisfaction, confidence and trust in doctor, care planning). We report the results of the regression as standardized regression coefficients. Only those practices with at least 100 GPPS respondents, at least 1 clinical domain with at least 100 patients, and with no missing data for practice population demographics or practice medical team characteristics were included.

Shrunken estimates were used to reduce the effects of measurement error by adjusting practice scores toward the overall mean according to the amount of measurement error, and therefore obtaining more accurate estimates of the underlying association of clinical quality and patient experience at the practice level.[8] Similarly, patient-mix adjustment removes effects of patient characteristics on measured quality and thus improves the accuracy of the estimates of underlying 2 dimensions of quality and their correlation.[9,10] Shrunken case-mix adjusted estimates of the GPPS scores were obtained using SAS v9.2 (SAS Institute). All other analysis was performed using Stata v11.1 (StataCorp LP).

References

1. Ipsos MORI. *The GP Patient Survey*. http://www.gp-patient.co.uk.
2. Campbell J, Smith P, Nissen S, Bower P, Elliott M, Roland M. 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.
3. *The Technical Annex for the GP Patient Survey Annual Report 2009-2010*. London: Ipsos Mori; 2010. http://www.gppatient.co.uk/.../Y4Q4_AnnualTechnical_unweighted.pdf.
4. Noble M, McLennan D, Wilkinson K, Whitworth A, Barnes H, Dibben C. *The English Indices of Deprivation. Department for Communities and Local Government.* 2007.
5. http://www.communities.gov.uk/publications/communities/indiciesdeprivation07.
6. Roland M. Linking physician pay to quality of care: a major experiment in the UK. *New Engl J Med*. 2004; 351(14):1448-1454.
7. Doran T, Fullwood C, Reeves D, Gravelle H, Roland M. Exclusion of patients from pay-for-performance targets by English physicians. *N Engl J Med*. 2008;359:274-284.
8. Health and Social Care Information Centre. http://www.ic.nhs.uk.
9. Adams JL, Mehrotra A, Thomas JW, McGlynn EA. Physician cost profiling – reliability and risk of misclassification. *N Engl J Med*. 2010;362:1014-1021.
10. 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;44:501-518.
11. Zaslavsky, AM. Epstein AM. How patients' sociodemographic characteristics affect comparisons of competing health plans in California on HEDIS quality measures. *Int J Qual Health Care.* 2005;17(1):67-74.