Physician Profiling: Methodologic Safeguards Needed

Greenfield S, Kaplan SH, Kahn R, et al. Profiling care provided by different groups of physicians: effects of patient case-mix (bias) and physician-level clustering on quality assessment results. Ann Intern Med 2002;136:111–21.

Study Overview

Objective. To investigate the effects of patient case-mix and physician-level clustering on differences in quality of diabetes care between endocrinologists and generalists.

Design. Retrospective medical record review that collected clinical data for 1 year prior to the recruitment date and a cross-sectional survey that collected patient case-mix variables and patient satisfaction scores.

Setting and participants. 29 solo and group practices were selected from diverse regions in the United States; 15 were endocrinology sites and 14 were generalist sites. There was an average of 67 patients sampled per site, with 90% (26/29) of the sites recruiting at least 27 patients. Only those patients who received most of their diabetes care at the study site were included.

Methods. Regression models were used to calculate estimated odds ratios (ORs) for each process and outcome measure. ORs were calculated with no adjustment, adjustment for patient case-mix only, adjustment for physician-level clustering only, and adjustment for patient case-mix and physician-level clustering. The models were adjusted for patient age, ethnic minority status, gender, education, diabetes duration, health status, and insulin treatment. The investigators used generalized estimation equations to investigate the effect of physician-level clustering on the outcomes.

Main outcome measures. Process and outcomes measures collected included lipid profile, dilated eye examination, foot examination, blood pressure, any test for urinary protein, hemoglobin A_{1c}, any documentation of self-monitoring of blood glucose (SMBG), and patient satisfaction. For hemoglobin A_{1c}, the investigators used only the most recent recorded value. Satisfaction was measured using a 5-point Likert scale with the proportion of "excellent" responses being used for the analysis.

Main results. Of the 1750 patients participating in the study, 52% were women and 81% were white. In the unadjusted patient-level analysis, there were statistically significant differences between endocrinologists' patients and generalists' patients in several domains. Endocrinologists performed better in the obtaining of hemoglobin A_{1c} measurements, urine protein measurements, eye and foot examinations, and SMBG. The specialists also had better outcomes for hemoglobin A_{1c} values, high-density lipoprotein cholesterol values, blood pressure, and patient satisfaction. Additionally, endocrinologists were found to have a case-mix of patients who were younger, better educated, had a longer duration of diabetes, and were more likely to be taking insulin. After adjusting for this patient case-mix, the difference between the performance of lipid tests, eye examinations, and SMBG was no longer significant. After further adjusting for physician-level clustering, only patient satisfaction remained significant, with endocrinologists having more satisfied patients (OR, 1.95; 95% confidence interval, 1.14–3.33)

Conclusion. Profiling physicians to compare quality of care may not be accurate unless careful adjustments are made to account for patient case-mix and physician-level clustering.

Commentary

Patients are demanding high-quality health care [1]. In order to make educated decisions concerning health care, many patients want access to information on quality [2]. One metric for presenting quality information is through physician profiling, the comparison of different physicians (individuals or groups) with regards to a specific process or outcome measurement. In this well-designed study, Greenfield and colleagues have identified a major design weakness in profiling studies. By using diabetes care as an example, the authors dramatically demonstrated how one can be misled by the results of a physician-group comparison if certain methodologic safeguards are not in place. Because patients seen by an individual provider are often more similar to each other than patients seen by another provider, the statistical assumption that all individuals within one provider's practice are independent is threatened. This physician-level "clustering" can have a dramatic effect on the power of a study to detect a difference between 2 groups. To account for this loss of power, physician-profiling studies must have sufficient numbers of physicians and patients per physician in order to truly detect a difference. Unfortunately, adjustments are rarely made for patient case-mix and physician-level clustering.

There were some minor limitations to the study. Because physicians volunteered to enter the study and were then

(continue on page 121)

(continued from page 118)

responsible for deciding which patients to recruit, some selection bias could have occurred. However, as the authors discerned, the variation between the individual physicians was so large that enrolling more patients was unlikely to influence their results. An additional concern is that process and outcome measures were collected from retrospective chart review. However, it is unlikely that the groups had markedly different documentation styles.

Applications for Clinical Practice

Careful attention must be paid to studies designed to compare quality measures between physician groups. Unless the sample size is adequately large and patient case-mix and physician-level confounding are taken into account, the results could be inaccurate.

–Review by Harvey J. Murff, MD

References

1. Eisenberg JM. Measuring quality: are we ready to compare the quality of care among physician groups? Ann Intern Med 2002;136:153–4.
2. Hibbard JH, Harris-Kojetin L, Mullin P, et al. Increasing the impact of health plan report cards by addressing consumers' concerns. Health Aff (Millwood) 2000;19:138–43.