

Systems Learning from Physician Performance Data

By Patricia H Parkerton, PhD, MPH
Gary A Feldbau, MD; Hugh L Straley, MD

Abstract

Context: Data collected in 1998 on primary physician performance, including Health Plan Employer Data and Information Set (HEDIS) measures, were the basis of reports distributed quarterly to 194 primary care physicians at 25 medical centers in Group Health Cooperative. Here, we summarize results of research designed to assess reliability of measures of physician performance and to identify practice components which influence patient outcome. Various aspects of these results are published in *Medical Care*, *The Journal of General Internal Medicine*, and *Family Medicine*.

Design: Summary of results from studies that used retrospective analysis of administrative data on physician performance measures and practice structures.

Main Outcome Measures: Twenty-three HEDIS measures of physician performance, both individual and grouped into aggregate measures: cancer screening, diabetic management, patient satisfaction, and ambulatory costs.

Analysis: Bivariate and sequential sets of multiple regression models controlled for selected patient panel and physician characteristics.

Results: Although individual HEDIS measures were reliable when used to assess physician performance, aggregated measures were more reliable. Physician continuity was not associated with patient outcome, but practice coordination (measured by shared practice, years of team tenure, and medical clinic size) was significantly associated with improvement in cancer screening, diabetic management, and patient satisfaction. Performance assessment of physicians with reduced appointment hours or part-time status was associated with improved cancer screening and diabetic management.

Conclusions: Assessing physician performance data on individuals yielded useful collective clinical practice information. Analyzing physician performance data collectively can identify effective primary care practice structures and processes and benefit patient care.

Introduction

The Group Health Cooperative of Puget Sound (Group Health) comprises medical centers and networked physician practices in the State of Washington and selected counties in northern Idaho; Group Health Permanente (GHP) physicians work in these medical centers. From 1997 through 1998, GHP primary care physicians received quarterly reports of data collected on their performance. The physicians were working toward improved practice. However, results from practice and performance measures continued to vary widely within and between the 25 medical centers. The Associate Medical Director for Quality and Research and Director of the Sandy MacColl Institute for Health Care Improvement agreed to sponsor a doctoral student to assess the data and its value to improving primary care practice. The objectives were to evaluate reliability of current performance assessment measures and to determine if aggregation of these measures of physician performance was appropriate, to identify components or structures of physician practice that influence patient outcomes, and to extract ideas for practice improvement from the results of this research.

The collaboration yielded new insights into physician performance assessment and practice structures, and articles on different aspects of the research were published in *Medical Care*, *The Journal of General Internal Medicine*, and *Family Practice*.¹⁻³ Here we summarize the published research and results.

Methods

Interviews with 30 key physicians and administrators generated the following research questions, which guided the type of data collected and the data analyses.

1. Are selected Health Plan Employer Data and Information Set (HEDIS) measures, which were developed to assess health plans, reliable when used to assess primary care physician performance?
2. Are cancer screening, diabetic management, patient

The collaboration yielded new insights into physician performance assessment and practice structures ...



Patricia H Parkerton, PhD, MPH, (top) is an Assistant Professor of Health Services in the UCLA School of Public Health. She previously was Manager of Operations for Kaiser Foundation Health Plan of Connecticut. E-mail: parkert@ucla.edu.
Gary A Feldbau, MD, (left) is CMO and Senior VP at Community Health Plan of Washington and Community Health Network of Washington. E-mail: feldga1@msn.com.
Hugh L Straley, MD, (right) is the Medical Director of Group Health Cooperative and President/Board Chair of Group Health Permanente. E-mail: straley.h@ghc.org.

satisfaction, and ambulatory costs reasonable and reliable aggregate measures of physician performance?

3. If primary care physicians spend fewer hours in direct patient care, do patient outcomes—particularly patient satisfaction—suffer?
4. Does continuity of care with the primary care physician influence patient outcome?
5. Are specific practice structures beneficial to patient outcome?

Setting, Study Design, and Data Collected

The study population was all 194 GHP family practitioners and general internists who provided ambulatory primary care services for at least nine months during 1998 to a designated patient panel from 320,000 adult Group Health members at 25 medical centers in western Washington. Physicians who provided urgent care exclusively were excluded. The centers varied in size and complexity, but all provided primary care, radiology, laboratory, pharmacy, and business services, and some provided specialty services.

The cross-sectional research design used existing administrative data integrated with additional practice and physician data that we collected. Quarterly reports distributed by the medical group to the physicians for two years before this study included measures of individual physician performance, appointment access, panel size and composition, and patient case-mix. Additional data included physicians' board certification, validation for specialty, and gender (all obtained from medical directories and from the American Medical Association Web site),^{4,5} and practice structure data gathered from the Human Resources department and practice leaders. These data were merged with the performance data using a random identifier to protect physician confidentiality.

This research was approved by the Institutional Review Board associated with Group Health and its research center. Funding sources placed no constraints upon this research, and Group Health allowed the researcher access to the organization and its data without determining the topic of inquiry, its analysis, or interpretation.

Physician Performance Measures

Performance measures that are systematically collected at the physician level are also considered as outcomes for the physicians' patient panels. Selected measures were grouped to form the following four aggregate measures of performance: diabetic management, cancer screening for women, patient satisfaction, and ambulatory costs

(Table 1).^{1,2} Ambulatory cost measures were averaged for the year, and other component performance data were reported as a rolling average of the previous year.

"Higher rates correspond to better outcomes for cancer screening, diabetic management compliance, and patient satisfaction, whereas the preferred cost outcome is lower. These measures are widely used,⁶ have sufficient patient populations to provide reliable assessment,⁷ and represent different aspects of care. For cancer screening, which combines rates of screening [in] different subpopulations, the aggregate measure is the mean of component measures,

Table 1. Performance measures: Aggregates and their component measures (n = 194)

Aggregate and component measures	Result (range)
Patient satisfaction^a	Percentage "excellent" responses
Friendliness and caring	14-80
Attention paid	20-79
Opportunity to ask questions	20-71
Explanations given about care	10-68
Support on ways to stay healthy	14-63
Time spent	10-54
Thoroughness and competence	17-76
Aggregate mean, 42%	17-67
Diabetic management	Percentage tested
Annual foot examination	11-100
Annual retinal examination	36-90
Microalbuminuria testing	50-100
Hemoglobin A _{1c} testing	68-100
Aggregate mean, 81%	54-99
Cancer screening for women	Percentage tested
Mammography, age 52-64 years	53-92
Papanicolaou test, age 21-64 years	53-88
Aggregate mean, 76%	62-87
Ambulatory costs^b	\$ per member per month
Primary care, Medicare	36-78
Primary care, non-Medicare	21-46
Special care, Medicare	37-145
Special care, non-Medicare	11-39
Radiology, Medicare	2-20
Radiology, non-Medicare	2-8
Laboratory, Medicare	3-12
Laboratory, non-Medicare	1-6
Pharmacy, Medicare	22-87
Pharmacy, non-Medicare	11-56
Aggregate mean, \$85	62-128

^a Measured on patient survey.

^b n = 176

Adapted and reproduced with permission of the publisher from: Parkerton PH, Smith DG, Belin TR, Feldbau GA. Physician performance assessment: nonequivalence of primary care measures. Med Care 2003 Sep;41(9):1034-47¹ and Parkerton PH, Smith DG, Straley HL. Primary care practice coordination versus physician continuity. Fam Med 2004 Jan;36(1):15-21.²

both of which summarize comparable screening rates for subgroups of each patient panel. Assuming a direct relationship between physician services and patient outcomes, these aggregate measures are indicators of both physician performance and the cumulative quality of care received by their patients. The measures of clinical processes, which included preventive services and disease treatment, and outcome measures, which included patient satisfaction, are included

in HEDIS measures required of managed care organizations. Furthermore, the cost measure is consistent with National Committee on Quality Assurance report card requirements.^{16,17}

Independent Predictors of Patient Outcomes

The independent predictors of physician performance were hours of direct patient care, continuity of care, and practice structure. Physicians were considered full-time by Group Health Permanente if they had ten appointment sessions (35 hours) scheduled for patient appointments each week; physicians with fewer sessions were considered part-time. Three sessions per week were the fewest a primary care physician could work and have a designated patient panel and therefore be included in this research. Physician direct patient care hours were determined by scheduled appointment hours, which ranged from 10 to 35 hours per week. Clinician continuity was determined by the percentage of primary care physician visits to total physician visits in one year. Practice coordination was assessed through the following three practice structures: shared practice, where two or three physicians accept joint responsibility for patients; clinical team tenure, defined as the number of years each physician worked with most physicians in the team; and medical clinic size, defined by number of physicians.

Statistical Analyses

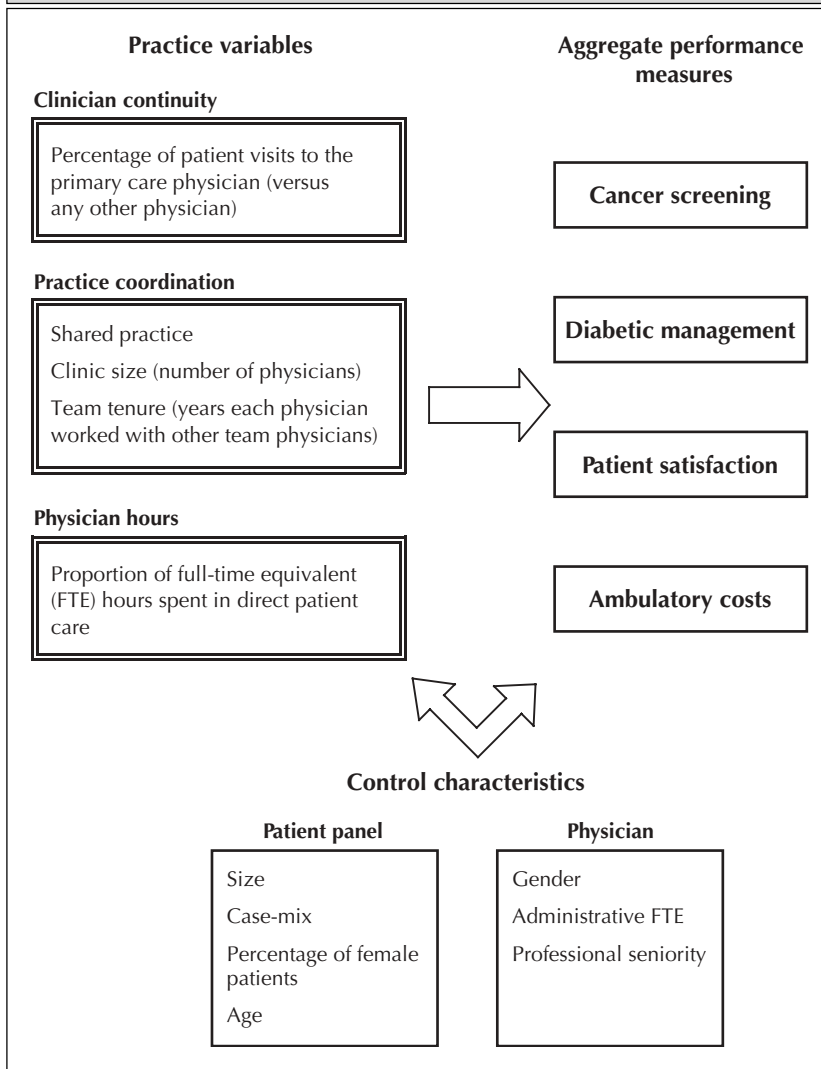
Sequences of ordinary-least-squares regression equations and comparisons by rank were used to assess physician performance and to determine practice influences on performance. Figure 1 shows the model of practice variables assessed for influence on physician performance measures as well as patient and physician characteristics controlled for in the analysis. The physician performance (patient outcome) analysis controlled for characteristics of physicians (administrative role, gender, seniority) and their patient panels (size, chronic disease score,^{8,9} gender, and age). Of the seven characteristics of physicians and their patient panels, five were significantly associated with one or two patient outcomes and were therefore necessary control variables.

Results

Wide Practice Variation for Each Performance Measure

The data had sufficient variation and number of cases to yield statistically significant results. Figure 2 shows the variation in aggregated measures of performance:

Figure 1. Model of the statistical analysis used to determine influence of practice variables on aggregated measures of physician performance



Patient panel size calculated by dividing the number of patients in each physician's panel by the full-time equivalent (FTE) hours the physician is scheduled for clinical appointments. 1 FTE = 35 hours per week. Case mix calculated by using the chronic disease score (determined by prescription data), percentage of females in patient panel, and mean age of patients in panel.

The percentage of patient satisfaction surveys returned with a response of “excellent” ranged from 17% to 67%; the percentage of patients who received appropriate tests for management of diabetes ranged from 54% to 99%; the percentage of eligible women who were screened for cancer ranged from 62% to 87%; and ambulatory costs per member per month ranged from \$62 to \$128.

Physicians’ Performance Varies by Measure

Are selected HEDIS measures, which were developed to assess health plans, reliable when used to assess primary care physician performance? Perhaps—if the physician’s practice is large enough (eg, contains enough diabetic patients¹⁰) and the measures are used independently.¹

Are cancer screening, diabetic management, patient satisfaction, and ambulatory costs reasonable and reliable aggregate measures of physician performance? Yes, when the aggregate measures are each evaluated, and performance is evaluated for groups of physicians rather than individual physician performance.

Each aggregate measure was reliable and independent, but loosely predictive, of the others. Each aggregate measure was significantly correlated with one or two of the other measures; high cancer screening rates correlated with close diabetic management and with high patient satisfaction scores; high diabetic management rates correlated with high cancer screening rates; and high patient satisfaction scores correlated with high cancer screening rates and high ambulatory costs. Further, physician performance was inconsistent across aggregate measures. More than 70% of the physicians ranked in the top third for at least one measure, but 80% of these same physicians ranked in the lowest third for a different measure. Sixty percent of the physicians ranked in the top third for one measure and in the bottom third for another.¹

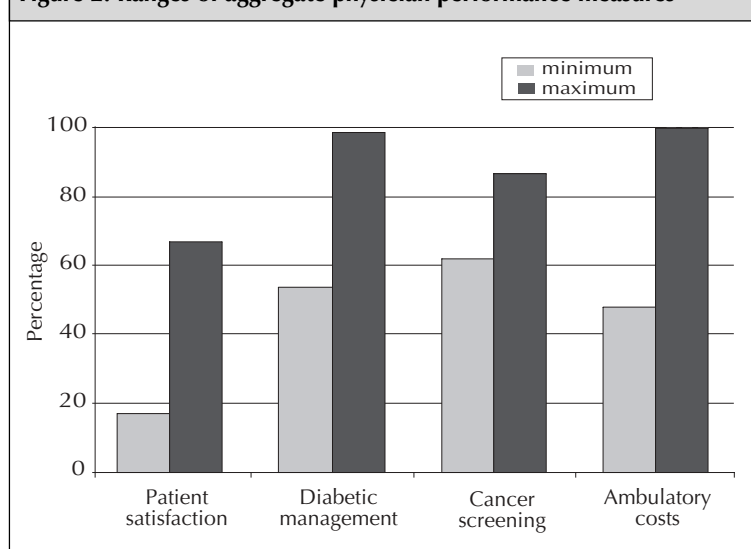
“Assessments of individual physicians with current performance measures may identify areas in which improvement is needed and facilitate provision of feedback to improve performance quality and efficiency. However, these performance measures, singly or as a unit, should be used cautiously to select, motivate, and reward physicians, or to encourage consumer assessment. There are relationships among physician performances in cancer screening, diabetes management, patient satisfaction and ambulatory costs. However these relationships are inconsistent across all physicians and unreliable for individual physicians.”¹

Part-time Practice Performance Not Worse

If primary care physicians spend fewer hours in direct patient care, are reduced hours associated with reduced patient outcomes—particularly patient satisfaction? No. Of the 194 physicians, 39% were considered employed full time by GHP, but because of other administrative duties, 85% of these full-time physicians worked less than full time in direct patient care (Figure 3). Physicians’ direct patient care hours showed a bimodal distribution: 4% worked less than half time (the commitment required for fringe benefits); 30% worked half time; and 20% worked full time in direct patient care.³

After adjusting for potential confounders, our analysis showed that as physician direct patient care hours decreased by 10%, the rate of cancer screening for women increased by 0.7% ($p = .010$), and the rate of diabetic management increased by 1.1% ($p = .008$). No association existed between physician direct patient care hours and patient satisfaction ($p = .212$) or ambulatory costs ($p = .323$). Although the data supported the analysis of continuous data at a minimum of three FTE (ten hours), no “threshold” of performance was found.³

Figure 2. Ranges of aggregate physician performance measures



Diabetic management and cancer screening results are percentage of patients tested. Patient satisfaction is percentage of responses of “excellent” on survey. Ambulatory cost minimum is calculated as percentage of maximum.

Adapted and reproduced with permission of the publisher from: Parkerton PH, Smith DG, Belin TA, Feldbau GA. Physician performance assessment: nonequivalence of primary care measures. *Med Care* 2003 Sep;41(9):1034-47.¹

Physician Continuity Not Related to Outcomes

Does continuity of care with the primary care physician influence patient outcomes? No. Physician continuity was not significantly associated with any patient outcome measure.²

Practice Structures Coordinate and Improve Care

Are specific practice structures beneficial to patient outcome? Yes. Each of the three practice structures was positively associated with some performance measures, and team tenure was strongly associated with all four outcome measures.² Aspects of practice coordination, as represented by three practice structures (shared practice, medical clinic size, and team tenure), were significantly and positively associated with cancer screening, diabetic management, or patient satisfaction but were not associated with ambulatory costs. Patient and physician characteristics had a large impact on costs. Both shared practice and larger medical clinic size were associated with a higher rate of cancer screening ($p < .001$) and with better diabetic management ($p < .01$). Physicians in shared practices were 7% more likely to screen patients for cancer and to better manage diabetic patients. No practice coordination variable,

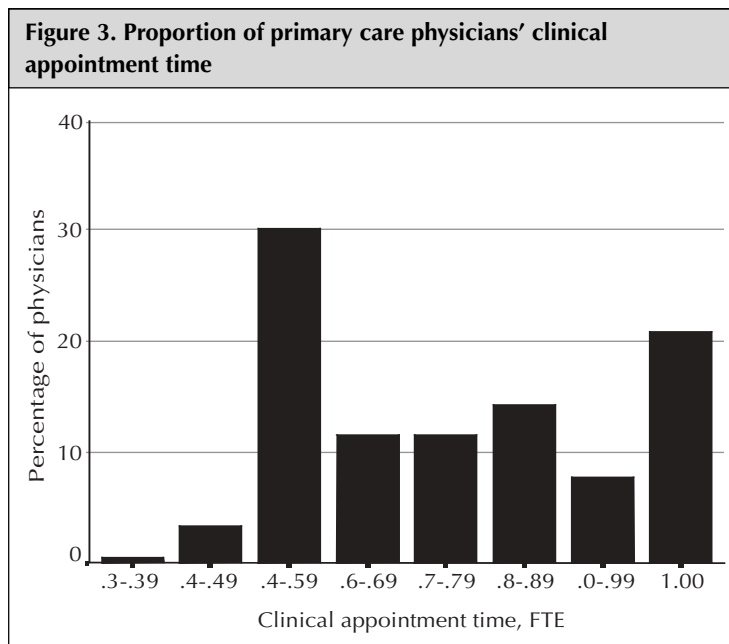
however, was significantly related to either patient satisfaction rating or to ambulatory costs.²

Discussion

Two of our objectives involved testing HEDIS measures for reliability when used to assess physician performance. We found that each individual measure was reliable over time and that reliability increased when individual measures were grouped with correlated measures into four aggregate measures. However, physician performance was not consistent across measures, a finding that was also reported recently in Boston area clinics.¹⁰ Although some specific measures are significantly related, overall predictive value of any single measure is low. The aggregate performance measures appear to assess different aspects of practice; therefore, blending their results may mislead to conclusions. "Because these aggregate measures are not strongly correlated, an overall measure, or using [one] as a proxy for all, is not recommended. Care should be taken in assessing physicians based on narrow performance measures resulting from current inconsistency in performance and the evolution of quality measures."¹¹

Our efforts to identify influences on patient outcomes associated with physician practice organization were productive, although the results were not as anticipated. We found that part-time physicians performed as well or better on the aggregate measures—including the measure of patient satisfaction—than those who worked more hours. As appointment hours decreased, performance either held constant or improved. Contrary to expectations, the trend toward reduced clinical hours merited attention but was not a current problem. Moreover, physician continuity, which we encourage, did not reach higher levels of physician performance. In fact, patient satisfaction declined as continuity increased. We believe that the explanation for this lies in coordinating structures around the physician-patient communication mechanisms other than the traditional visit (eg, telephone, e-mail, team members). Each of three practice structures were positively associated with some patient outcome measures: shared practice, larger medical centers, and clinical team tenure of 4 to 15 years. Because these practice structures were selected for availability (ie, convenience), they may not be the most influential structures.

Instead of having lower performance results, primary care physicians who worked fewer direct patient care hours had slightly higher cancer screening rates, better diabetic management rates, and similar patient satisfaction scores and ambulatory costs compared with



n = 194. 1 FTE = 35 hours.

Adapted and reproduced with permission of the publisher from: Parkerton PH, Wagner EH, Smith DG, Straley HL. Effect of part-time practice on patient outcomes. *J Gen Intern Med* 2003 Sep;18(9):717-24.³

those of full-time physicians. Identifying the practice organization mechanisms used by physicians who work fewer hours may be a step toward improving outcomes for all patients.³ We found that physician continuity was not associated with better outcomes but that specific practice structures were. The widespread assumption that physician continuity is central to quality makes the absence of a positive relation with performance measures surprising and emphasizes the need to pursue other means of coordinating patient care services.²

Our final objective was to identify ideas for practice improvement. Although we considered the composition and roles within primary care teams to be potentially influential, gathering team data was difficult. We attempted to assess nursing and team pharmacist roles, but no administrative source was available, and our data were incomplete and lacked sufficient power to show statistically significant effects. Better understanding of team member roles and their optimization would be useful.

Limitations of this research include its focus on physicians working in a medical group within a health maintenance organization, in teams with other practitioners, and within a single organization. Generalizability of these findings is also limited by the reduced reliability of performance measures for physicians with smaller patient panels or who provide care, for example, for few diabetic patients.⁷ In addition, the physicians included in our research cared for patients who had comprehensive health insurance benefits, and the physicians functioned as gatekeepers to specialty services. Therefore, our analyses implicitly controlled for specialty, organization, health benefits, payment, access to service, and designation of primary physician.

Conclusions

Reduced physician hours and physician continuity did not reduce the four aggregate measures of patient outcome, and some primary care practice structures (shared practice, larger medical centers, clinical team tenure of 4 to 15 years) benefited patient outcomes.

Interviews with key leaders helped us to formulate useful research questions and to increase access to data. Individual primary care physician performance data yielded collective clinical practice information. Our analyses led to conclusions which differed from popular opinion and thus redirected some planning efforts.

Analyzing physician performance data can help us to identify effective primary care practice structures and processes and can ultimately benefit patient care. How-

ever, constructive use of physician performance data requires acknowledgment of both positive and negative performance by individual physicians so that accolades are supported and poor practices are not masked. Performance data on a population of physicians can also allow their efforts to be tracked to improve the practice environment. ♦

Acknowledgments

This research received financial support from The Blue Cross Blue Shield of Michigan Foundation. The Sandy MacColl Institute for Healthcare Improvement and Group Health provided resource support for data collection. Dean Smith, PhD, Edward H Wagner, MD, MPH; and Mary Richardson, PhD; provided academic guidance. Staff of Group Health providing insight and data access included Brian Austin; Susan Crissman, RN, MEd, MPH; Matthew Handley, MD; Michael Wanderer, MD; and Nirmala Sandhu, MPH.

Dr Parkerton received support for her doctoral program as an Agency for Healthcare Research and Quality (AHRQ) Fellow.

References

1. Parkerton PH, Smith DG, Belin TR, Feldbau GA. Physician performance assessment: nonequivalence of primary care measures. *Med Care* 2003 Sep;41(9):1034-47.
2. Parkerton PH, Smith DG, Straley HL. Primary care practice coordination versus physician continuity. *Fam Med* 2004 Jan;36(1):15-21.
3. Parkerton PH, Wagner EH, Smith DG, Straley HL. Effect of part-time practice on patient outcomes. *J Gen Intern Med* 2003 Sep;18(9):717-24.
4. American Medical Association. Directory of physicians in the United States. 36th ed. Chicago: American Medical Association, Division of Survey and Data Resources; 1999.
5. American Medical Association. Doctor Finder [Web site]. Available from: <http://dbapps.ama-assn.org/aps/amahg.htm> (Accessed July 18, 1999).
6. National Committee on Quality Assurance. HEDIS® 2000. List of measures. Washington (DC): National Committee on Quality Assurance; 2000. Available from: www.ncqa.org/Programs/HEDIS/index.htm (accessed March 21, 2000).
7. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *JAMA* 1999 Jun 9;281(22):2098-105.
8. Von Korff M, Wagner EH, Saunders K. A chronic disease score from automated pharmacy data. *J Clin Epidemiol* 1992 Feb;45(2):197-203.
9. Putnam K, Buist D, Fishman P. Chronic disease score as a predictor of hospitalization. *Epidemiology* 2002 May;13(3):340-6.
10. Gandhi TK, Francis EC, Puopolo AL, Burstin HR, Hass JS, Brennan TA. Inconsistent report cards: assessing the comparability of various measures of the quality of ambulatory care. *Med Care* 2002 Feb;40(2):155-65

Analyzing physician performance data can help us to identify effective primary care practice structures and processes and can ultimately benefit patient care.