

Variations of Physician Group Profiling Indicators for Asthma Care

I-Chan Huang, PhD; Gregory B. Diette, MD, MHS; Francesca Dominici, PhD; Constantine Frangakis, PhD; and Albert W. Wu, MD, MPH

Objective: To determine how much of the variation in physician group profiling for asthma care can be attributed to physician groups and how reliable those profiling indicators are.

Study Design: Cross-sectional study. Variations attributable to physician groups are presented using the intraclass correlation coefficient (ICC). The reliability of profiling results was determined using the ICC and sample size of the physician group.

Participants and Settings: Between July 1998 and February 1999, patients with asthma from 20 California physician groups were randomly selected to be surveyed; 2515 patients responded.

Main Outcome Measures: Quality indicators for physician group profiling were (1) National Asthma Education and Prevention Program guideline-based processes of care, including accessibility of asthma care, self-management knowledge about asthma care, use of inhaled bronchodilators, and use of inhaled corticosteroids, and (2) patient outcomes, including satisfaction with asthma care, improvement in health status, and emergency department visits and hospitalizations attributable to asthma.

Results: The variations attributable to physician group were small (< 10%) for process and outcome indicators. For process indicators, self-management knowledge had the highest ICC (9.83%), and use of inhaled bronchodilators had the lowest ICC (3.08%). For outcome indicators, satisfaction with asthma care had the highest ICC (9.53%), and hospitalization had the lowest ICC (1.35%). Despite low ICCs, a large sample size per physician group (n = 126) yielded acceptable reliability (≥ 0.80) for most profiling results.

Conclusions: The selected indicators for profiling asthma care at the physician group level were generally reliable. Sampling a sufficient number of cases is key to achieving useful results from profiling.

(*Am J Manag Care.* 2004;10:38-44)

Asthma is a common disease, characterized by inflammation of airways and reversible obstruction to airflow, that affects an estimated 14.6 million persons in the United States.¹ In response to repeated demonstrations of suboptimal asthma treatment, the National Asthma Education and Prevention Program (NAEPP) Expert Panel published *Guidelines for the Diagnosis and Management of Asthma* in 1991.² The guidelines, which were revised in 1997 and 2002, emphasize the importance of patient education and appropriate use of medications.^{3,4}

Leading quality oversight organizations assess performance of asthma care by individual physicians, physician groups, and health plans using NAEPP-based guidelines.⁵⁻⁷ The expectation is that provider profiling can increase provider accountability to improve quality of care, help to control healthcare costs, and guide consumers to high-quality providers.⁸ To date, most profiling indicators have been selected based on "clinical importance" and represent important processes or outcomes of care. However, little is known about whether actual variations in these indicators are large enough to discriminate among different providers. The amount of variation in provider profiles that can be truly attributed to providers, after adjusting for patient case mix, can be estimated using the intraclass correlation coefficient (ICC).⁹⁻¹² For judging provider performance, it is useful to have profiling indicators with a high ICC, implying that indicator scores tend to be similar for patients cared for by the same provider, and providing larger differences across providers.

However, the ICC only provides the percentage of variations due to provider effects and therefore cannot be the sole indicator of performance. When the number of patients per practice is small, even if the ICC is high, it is possible that results of profiling may be uninformative. Therefore, it is useful to apply a second indicator, the reliability of profiling, which considers the ICC together with the number of patients sampled from a provider.^{10,11}

There have been few studies of the ICC of profiling indicators. The range of variations attributable to providers varies depending on disease and the selection

From the Departments of Health Policy and Management (I-CH, AWW), Epidemiology (GBD, AWW), and Biostatistics (FD, CF), Bloomberg School of Public Health, and the Department of Medicine, School of Medicine (GBD, AWW), The Johns Hopkins University, Baltimore, Md.

This study was funded by the Pacific Business Group on Health, San Francisco, Calif. This study was presented in part at the 20th Annual Research Meeting of AcademyHealth; Nashville, Tenn; June 27-29, 2003.

Address correspondence to: Albert W. Wu, MD, MPH, Department of Health Policy and Management, Bloomberg School of Public Health, The Johns Hopkins University, 624 North Broadway, Baltimore, MD 21205-1901. E-mail: awu@jhsph.edu.

of an indicator. In general, the variations attributable to providers are small (< 10%).^{10,13-17} For example, Hofer and colleagues¹⁰ assessed physician profiles for the care of patients with type 2 diabetes mellitus and found that the overall variance in hospitalization rates attributable to physician practice was only 1%. Krein et al¹⁶ found that the ICCs of diabetes process and outcome indicators at the primary care provider level ranged from 0% to 9%. Sixma and colleagues¹³ showed that the ICC of patient satisfaction with general practitioners was 5% to 10%. A review by Campbell et al¹⁸ suggested that at the individual practice level the ICCs of process indicators were higher than those of outcome indicators.

Fewer studies^{10,17} have examined the reliability of profiling indicators. Hofer et al¹⁰ suggested that the reliability of physician profiles for hospitalization rates for patients with type 2 diabetes mellitus was only 0.17. Poor reliability of profiling results at the individual physician level can be due to small panels of patients.¹⁰

This study evaluated how much of the variance of physician group profiling is attributed to physician group effects and how reliable physician group profiling is for process and outcome indicators. If the variation attributable to physician groups and the reliability of profiling results are small, then current profiling practices may need reexamination. We used consistency with asthma guidelines and patient outcomes as performance indicators and the physician group as the unit for profiling.

METHODS

Study Setting

This study was conducted in conjunction with 20 California physician groups that participated in the 1998 Asthma Outcomes Survey. The Asthma Outcomes Survey was initiated by the Pacific Business Group on Health and by HealthNet to evaluate, improve, and report on the quality of asthma care at the physician group level.¹⁹

Although profiling often focuses on individual physicians or entire organizations, experts have suggested that profiling of physician groups may be useful. There is a practical need for managed care organizations to profile physician groups, as patients tend to select plans based on individual physicians or physician groups.^{20,21} Therefore, profiles at the physician group level could enhance consumer choice. Individual groups may be more receptive to data on their own practice rather than data from the entire organization. Physicians may want to join groups that they believe deliver better care, and managed care organizations want to contract with them.

Sample Selection and Data Collection

Details on sample selection and data collection have been described.¹⁹ Briefly, the 20 participating physician groups used administrative materials to identify all managed care patients with at least 1 asthma-related encounter in the outpatient, emergency, or inpatient setting (identified by *International Classification of Diseases, Ninth Revision, Clinical Modification* code 493.xx) between January 1, 1997, and December 31, 1997. To reduce misclassification of chronic obstructive pulmonary disease as asthma, we restricted our subjects to those younger than 55 years. Patients had to be continuously enrolled in the physician group for that calendar year. From eligible patients, the study randomly selected a sample of 650 patients from each physician group. If a physician group had fewer than 650 eligible patients, then all eligible patients were sampled.

Patient data were collected by mailed survey. The instrument was largely based on the Health Survey for Asthma Patients developed at The Johns Hopkins University for the Outcomes Management System Consortium Asthma Project.²²⁻²⁴ The survey asked about patient characteristics, general health, asthma symptoms, effect of asthma on functioning, asthma medications and treatment, self-management knowledge and activities, access to care, and patient satisfaction. The survey was fielded between July 1998 and February 1999. A total of 2515 responses were obtained, for a response rate of 32.2%.

Performance Indicators

In this study, we selected processes of asthma care and patient outcomes as indicators for publicly reported physician group comparisons. We evaluated the score variability and reliability for those indicators.

Processes of care were assessed by consistency with the NAEPP asthma guidelines, including accessibility of asthma care, self-management knowledge about asthma care, use of inhaled bronchodilators, and use of inhaled corticosteroids. Access to asthma care included accessibility of clinicians by telephone, for appointments, and to get asthma medications. Self-management knowledge measured ability to manage asthma flares, appropriately adjust asthma medication, and identify asthma triggers. For medication use, the NAEPP guidelines advocate inhaled corticosteroids as the most consistently effective long-term control medication and recommend inhaled bronchodilators (or β_2 -agonists) as rescue medications.³ In the survey, patients were rated on the number of puffs of inhaled bronchodilators and inhaled corticosteroids used every day. We dichotomized responses for inhaled bronchodilator use into 8 puffs or fewer as "no overuse" and more than 8 puffs as "overuse," and inhaled corti-

costeroid use into 4 puffs or fewer as “underuse” and more than 4 puffs as “no underuse,” based on guidelines and recommendations before the NAEPP guidelines were updated in 2002.²³

Outcome measures included satisfaction with asthma care during the past week, improvement in health status during the past week, and emergency department visits and hospitalizations attributable to asthma during the past year. We dichotomized responses on patient satisfaction into “greater satisfaction (excellent or very good)” vs “less satisfaction (good, poor, or fair),” improvement in health status into “greater improvement (much better or somewhat better)” vs “less improvement (about the same, somewhat worse, or much worse),” and emergency department visit and hospitalization into “no visit” vs “visits 1 or more times” and “no hospitalization” vs “hospitalizations 1 or more times.”

Risk Adjustment

Candidate risk-adjustment variables were collected from the patient survey, including patient age, sex, education level, type of health insurance, asthma severity, number of asthma-related comorbidities, and health status (Medical Outcomes Study 36-Item Short-Form Health Survey physical component score and mental component score). Asthma-related comorbidities included rhinitis, sinusitis, chronic bronchitis, heartburn (gastroesophageal reflux), emphysema, and congestive heart failure. The study measured asthma severity using responses to several questions to approximate the NAEPP’s 4 severity strata (mild intermittent, mild persistent, moderate persistent, and severe persistent).³ We measured severity using patients’ reports of the frequency of symptoms (cough, sputum, wheezing, chest tightness, and shortness of breath), the frequency of nocturnal symptoms, and the chronicity of symptoms between attacks. Severity was determined by the greatest severity in the responses to any of these questions.^{19,23}

Statistical Analysis

χ^2 and *t* tests were used to identify bivariate relationships between performance indicators and candidate risk-adjustment variables. We selected risk-adjustment variables that were statistically significant ($P < .05$) for inclusion in multivariate risk-adjustment models. We included all asthma patients to calculate the ICCs of profiling indicators. However, based on recommendations of the 1997 NAEPP guidelines, we only included asthma patients who had moderate persistent and severe persistent severity for the inhaled corticosteroid use indicator.³

We used Bayesian hierarchical modeling (HM) to quantify variations of performance indicators across the

20 physician groups that were attributed to physician groups. The use of Bayesian HM is regarded as a more appropriate approach than conventional approaches, as it takes into account the statistical uncertainty of each group-specific performance and the natural heterogeneity of the true group-specific performances, a key source of uncertainty of these analyses.²⁵ The major advantage of HM is that it allows us to assess physician group performance by quantifying random intercepts of logistic regressions at the patient level.^{25,26} Most important, HM can appropriately partition variations of performance measures across physician groups into between-physician group variability and within-physician group variability, and the variance estimates can then be used to produce “shrunk” estimates that are better estimates of the group effects.^{12,27} Estimates for groups with small case numbers are more likely to shrink toward the grand mean than those for groups with large case numbers.

The percentage of variability attributable to physician group effects relative to the overall residual variability can be estimated using the ICC as follows⁹⁻¹²:

$$\text{ICC} = \frac{\text{Variation Between Physician Groups}}{\text{Variation Between and Within Physician Groups}}$$

We adopted the method by Turner et al²⁸ to calculate ICCs for binary performance indicators. The estimations of ICCs under Bayesian HM were carried out using Markov chain Monte Carlo simulation. We used a uniform prior for the ICC estimation because thus far there is not much information available for the study of provider profiling. Markov chain Monte Carlo simulation comprised a burn-in of 500, followed by a further 5000 iterations, during which the posterior distribution of ICCs was monitored.²⁹

We calculated the reliability by combining the information of the ICC and the mean sample size (*n*) across the 20 physician groups using the following equation^{10,11}:

$$\text{Reliability} = \frac{n \times \text{ICC}}{1 + [(n-1) \times \text{ICC}]}$$

Based on the second equation, we can further calculate the required sample size based on the expected reliability of profiling results. To date, there is no agreed standard for judging the reliability of physician group profiling. Most research suggests that the reliability should be 0.80 or better.¹⁰ Another source suggests that acceptable reliability is at the level of 0.70.³⁰ Stata 7 (StataCorp LP, College Station, Tex) was used for bivariate analyses and WinBUGS 1.3²⁹ for ICC calculation.

RESULTS

Characteristics of Physician Groups and Respondents

Table 1 gives the characteristics of the participating physician groups. Ten physician groups (50.0%) were classified as medical groups, 7 (35.0%) were independent practice associations, and 3 (15.0%) were foundation or community clinics. All of the 20 were multispecialty groups. Table 2 gives the characteristics of the 2515 asthma patients included in the study. Patients ranged in age from 18 to 56 years (mean ± SD, 39.9 ± 9.5 years), 71.2% were female, 70.3% were white and 5.1% African American, and 81.6% had at least some college education. In terms of clinical characteristics, 14.4% had mild intermittent asthma, 19.2% had mild persistent asthma, 49.3% had moderate persistent asthma, and 17.1% had severe persistent asthma. The mean ± SD number of comorbidities was 2.1 ± 1.4.

Variations in Performance Indicators Attributable to Physician Groups

Table 3 gives the proportion, ICC, and reliability for each performance indicator. The ICCs using Bayesian HM indicated that, for process and outcome indicators, variations attributable to physician groups were small (< 10%). Indicators of guideline consistency demonstrated slightly higher ICCs than outcome indicators. Among indicators of guideline consistency, self-management knowledge about asthma care had the highest ICC (9.83%), while use of inhaled bronchodilators had the lowest ICC (3.08%). Among patient outcome indicators, satisfaction with asthma care had the highest ICC (9.53%), and hospitalization had the lowest ICC (1.35%).

Reliability of Profiling Results

Table 3 gives the reliability of profiling results. In general, the reliability of indicators at the physician group level was acceptable based on the criterion of 0.80 or better. The reliability ranged from 0.60 to 0.92. Indicators of consistency of care with guidelines demonstrated slightly greater reliability than outcome indicators. Self-management knowledge about asthma care had the highest reliability

(0.92), and patient satisfaction with asthma care had the second highest reliability (0.91). In contrast, inhaled bronchodilator use and hospitalization were less reliable, at 0.77 and 0.60, respectively.

Sample Size Needed to Achieve Reliable Profiling

The Figure demonstrates the relationship between the reliability of profiling and desired sample size. Given a fixed ICC, the relationship between the reliability of profiling and sample size per physician group was exponential. Indicators with lower ICCs usually require a larger sample size per group than the 126 per group for the present sample to achieve acceptable reliability. If we set the reliability at the levels of 0.70 and 0.80, larger sample sizes (170 and 292, respectively, per group) were needed only for the hospitalization indicator. If we assumed a stringent reliability level of 0.90, larger sample sizes were needed for indicators of accessibility, bronchodilator uses, emergency department visits, and hospitalization visits (159, 283, 216, and 658, respectively, per group).

DISCUSSION

Quality-of-care and performance oversight organizations are beginning to use consistency of care with asthma guidelines and patient outcomes as indicators

Table 1. Characteristics of Physician Groups*

Characteristic	Value
No. of asthma patients responding to survey per practice group	
1-100	35.0
101-200	55.0
≥ 201	10.0
Overall, mean (SD)	125.8 (56.0)
Location	
Northern California	40.0
Southern California	60.0
Operational status of practice group (at the end of 2003)	
Active	90.0
Closed	10.0
Type of physician group	
Medical	50.0
Independent practice association	35.0
Foundation or community clinic	15.0
No. of primary care physicians per practice group (n = 19)	
1-100	47.4
101-200	10.5
≥ 201	42.1
Overall, mean (SD)	201.9 (190.7)
Multispecialty practice group	100.0

*Data are given as percentages among 20 physician groups unless otherwise indicated.

Table 2. Characteristics of Patients With Asthma*

Characteristic	Value
Age, y	
18-24	7.2
25-34	22.0
35-44	34.6
45-54	33.2
≥ 55	3.1
Overall, mean (SD)	39.9 (9.5)
Female sex	71.2
Education	
≤ High school	18.4
College	65.3
Graduate	16.3
Health insurance status	
Private	69.1
Through employer	24.8
Through self-purchase	1.3
Public, Medicare or Medicaid	4.9
Other	96.5
Drug insurance coverage	
Asthma severity	
Mild intermittent	14.4
Mild persistent	19.2
Moderate persistent	49.3
Severe persistent	17.1
No. of comorbidities, mean (SD)	2.1 (1.4)
Selected comorbidity	
Sinusitis	38.0
Heartburn	31.2
Bronchitis	14.3
SF-36 score, mean (SD)	
Physical component	45.7 (10.3)
Mental component	47.4 (10.7)

SF-36 indicates Medical Outcomes Study 36-Item Short-Form Health Survey.
*Data are given as percentages among 2515 patients unless otherwise indicated.

for asthma care. Examining performance indicators of asthma care for 20 California physician groups, we found that variations attributable to physician groups (ie, ICCs) were small (< 10%). Among these indicators, attributable variation was larger for patient self-management of asthma care and satisfaction with asthma care. By contrast, the use of inhaled bronchodilators, emergency department visits, and hospitalizations showed less variations attributable to physician groups. We also demonstrated adequate reliability (≥ 0.80) of profiling indicators for asthma care at the physician group level, except for inhaled bronchodilator use and hospitalization.

Our findings have practical implications for managed care decision makers and organizations engaged in pro-

filed for asthma or other diseases. The results suggest that profiling indicators need to be selected carefully, based not only on clinical significance but also on higher ICCs and greater reliability. Variability in quality indicators is one of the attributes listed as desirable for Health Employer Data and Information Set performance measures. Variability is not the only desirable attribute for performance measures: for example, a measure that elicits uniformly low performance could still be useful to document room for improvement. However, it is important for groups interested in quality assessment to be aware that different indicators have different reliabilities. In considering potential indicators, the reliability can help to determine the feasibility of a given indicator. Our results suggest that self-management knowledge about asthma care and use of inhaled corticosteroids may be good process indicators for physician group profiling, and satisfaction with asthma care and improvement in health status may be useful outcome indicators. Profiles based on inhaled bronchodilator use and hospitalization are less useful.

Our results also reinforce that the reliability of profiling results at the physician group level is much higher than that at the individual physician level because of the larger sample sizes afforded. Using the physician group as the unit of profiling, the present sample size per group ($n = 126$) provided reliable profiling results (reliability, ≥ 0.80), except for the hospitalization indicator. Because the mean number of patients per individual physician was only 13, the reliability of each profiling indicator would be poor, ranging from 0.14 (hospitalization) to 0.54 (self-management knowledge). Hofer and colleagues¹⁰ showed that to achieve a reliability of 0.80 for profiling hospitalization and physician visit rates for diabetes care required at least 100 patients per physician group. In this regard, an indicator that requires a sample of approximately 300 patients to be reliable would only be collectable for the largest asthma practices and would generally not be relevant for individual providers.

In interpreting our findings, several potential limitations should be noted. First, there was a low response rate to the patient survey. The response rates, however, were similar across physician groups. The effect of a low response rate on comparisons across physician groups is important if scores used for profiling differ between the respondents and nonrespondents. Although we would have liked to compare the characteristics of respondents and nonrespondents across the 20 physician groups, we did not have patient charac-

teristics for nonrespondents. Second, the variation attributable to physician groups may be underestimated because of unmeasured confounders not included in our models. For example, we did not collect clinical assessments or non-patient characteristics, such as the supply of physician groups or hospitals in the market. Lack of adjustment for these factors may increase random variation.^{11,12} On the other hand, the variation attributable to physician groups may be overestimated because we cannot precisely partition overall variations into the physician group level. Based on clustering characteristics among patients, physicians, and physician groups, it is better to partition overall variations into 3 levels. However, the Pacific Business Group on Health project did not collect information for individual physicians, so we could not further partition the variation by individual physician and physician group. Finally, in this study, we intended to identify the differences among groups in which there were important variations attributable to group effects. However, we did not attempt to identify the effect of these variations on outcomes, which is another criterion to judge the usefulness of profiling exercises.

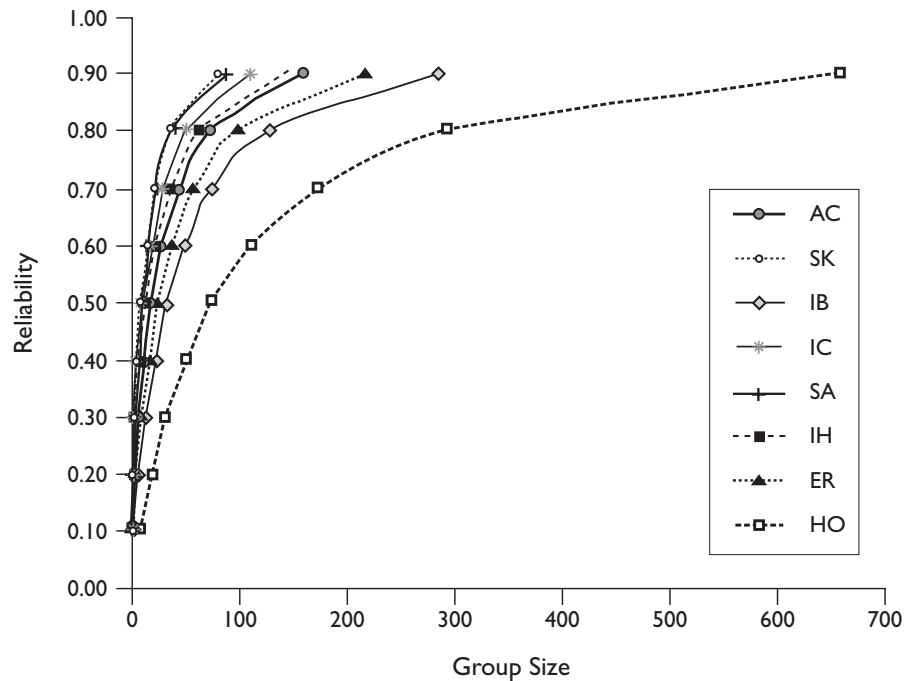
In conclusion, for performance profiling of asthma care across 20 physician groups, the variations attributable to physician groups were small. However, the

Table 3. Variations in Performance Indicators Attributable to Differences Among Group Practice

Performance Indicator	Proportion	Intraclass Correlation Coefficient (95% Confidence Interval), %	Reliability
Guideline-based processes of care consistency			
Accessibility of asthma care	0.70	5.37 (3.34-8.50)	0.85
Self-management knowledge	0.52	9.83 (6.47-14.64)	0.92
Bronchodilator inhaler, overuse	0.09	3.08 (1.71-5.22)	0.77
Corticosteroid inhaler, underuse*	0.73	7.55 (4.35-12.08)	0.89
Patient outcomes			
Satisfaction with asthma care	0.55	9.53 (6.14-14.28)	0.91
Improved health status	0.35	6.02 (3.75-9.33)	0.87
Emergency department visit due to asthma	0.22	4.00 (2.38-6.28)	0.81
Hospitalization due to asthma	0.06	1.35 (0.65-2.37)	0.60

*Asthma patients with moderate persistent or severe persistent severity.

Figure. Relationship Between Physician Group Sample Size and Reliability



AC indicates accessibility of asthma care (intraclass correlation coefficient [ICC], 0.05); SK, self-management knowledge (ICC, 0.10); IB, inhaled bronchodilators (ICC, 0.03); IC, inhaled corticosteroids (ICC, 0.08); SA, satisfaction with asthma care (ICC, 0.10); IH, improvement in health status (ICC, 0.06); ER, emergency department visit (ICC, 0.04); and HO, hospitalization (ICC, 0.01).

reliability of profiling results was generally acceptable because of sufficient case numbers at the physician group level. For profiling, we recommend the use of clinically important indicators with high reliability.

REFERENCES

1. **US Department of Health and Human Services.** *Action Against Asthma*. Bethesda, Md: US Dept of Health and Human Services; 2000.
2. **National Heart, Lung, and Blood Institute.** *National Asthma Education Program: Guidelines for the Diagnosis and Management of Asthma*. Bethesda, Md: US Dept of Health and Human Services; 1991.
3. **National Heart, Lung, and Blood Institute.** *Guidelines for the Diagnosis and Management of Asthma: Expert Panel Report No. 2*. Bethesda, Md: US Dept of Health and Human Services; 1997.
4. **National Heart, Lung, and Blood Institute.** *Expert Panel Report: Guidelines for the Diagnosis and Management of Asthma: Update on Selected Topics 2002*. Bethesda, Md: US Dept of Health and Human Services; 2002.
5. **Pacific Business Group on Health Web site.** Available at: <http://www.pbgh.org>. Accessed November 3, 2004.
6. **Foundation for Accountability Web site.** Available at: <http://www.facct.org/facct/site/facct/facct/home>. Accessed November 3, 2004.
7. **National Committee for Quality Assurance.** *The State of Health Care Quality: 2002*. Washington, DC: National Committee for Quality Assurance; 2002.
8. **Marshall MN, Shekelle PG, Leatherman S, Brook RH.** The public release of performance data: what do we expect to gain? A review of the evidence. *JAMA*. 2000;283:1866-1874.
9. **Gulliford MC, Ukoumunne OC, Chinn S.** Components of variance and intra-class correlations for the design of community-based surveys and intervention studies: data from the Health Survey for England 1994. *Am J Epidemiol*. 1999;149:876-883.
10. **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;281:2098-2105.
11. **Snijders TAB, Bosker RJ.** *Multilevel Analysis: An Introduction to Basic and Advanced Multilevel Modeling*. London, England: Sage Publications; 1999.
12. **Raudenbush SW, Bryk AS.** *Hierarchical Linear Models: Applications and Data Analysis Methods*. Thousand Oaks, Calif: Sage Publications; 2002.
13. **Sixma HJ, Spreuwenberg PM, van der Pasch MA.** Patient satisfaction with the general practitioner: a two-level analysis. *Med Care*. 1998;36:212-229.
14. **Davis P, Gribben B, Lay-Yee R, Scott A.** How much variation in clinical activity is there between general practitioners? a multi-level analysis of decision-making in primary care. *J Health Serv Res Policy*. 2002;7:202-208.
15. **Greenfield S, Kaplan SH, Kahn R, Ninomiya J, Griffith JL.** 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;136:111-121.
16. **Krein SL, Hofer TP, Kerr EA, Hayward RA.** Whom should we profile? examining diabetes care practice variation among primary care providers, provider groups, and health care facilities. *Health Serv Res*. 2002;37:1159-1180.
17. **Solomon LS, Zaslavsky AM, Landon BE, Cleary PD.** Variation in patient-reported quality among health care organizations. *Health Care Financ Rev*. 2002;23:85-100.
18. **Campbell M, Grimshaw J, Steen N.** Sample size calculations for cluster randomised trials: Changing Professional Practice in Europe Group (EU BIOMED II Concerted Action). *J Health Serv Res Policy*. 2000;5:12-16.
19. **Masland M, Wu AW, Diette GB, Dominici F, Skinner EA, Esquivel M.** *The 1998 Asthma Outcomes Survey*. San Francisco, Calif: Pacific Business Group on Health; 2000.
20. **Epstein AM.** Rolling down the runway: the challenges ahead for quality report cards. *JAMA*. 1998;279:1691-1696.
21. **Casalino LP, Devers KJ, Lake TK, Reed M, Stoddard JJ.** Benefits of and barriers to large medical group practice in the United States. *Arch Intern Med*. 2003;163:1958-1964.
22. **Steinwachs DM, Wu AW, Skinner EA.** How will outcomes management work? *Health Aff (Millwood)*. 1994;13(4):153-162.
23. **Diette GB, Wu AW, Skinner EA, et al.** Treatment patterns among adult patients with asthma: factors associated with overuse of inhaled β -agonists and underuse of inhaled corticosteroids. *Arch Intern Med*. 1999;159:2697-2704.
24. **Wu AW, Young Y, Skinner EA, et al.** Quality of care and outcomes of adults with asthma treated by specialists and generalists in managed care. *Arch Intern Med*. 2001;161:2554-2560.
25. **Christiansen CL, Morris CN.** Improving the statistical approach to health care provider profiling. *Ann Intern Med*. 1997;127:764-768.
26. **DeLong ER, Peterson ED, DeLong DM, Muhlbaier LH, Hackett S, Mark DB.** Comparing risk-adjustment methods for provider profiling. *Stat Med*. 1997;16:2645-2664.
27. **Cowen ME, Strawderman RL.** Quantifying the physician contribution to managed care pharmacy expenses: a random effects approach. *Med Care*. 2002;40:650-661.
28. **Turner RM, Omar RZ, Thompson SG.** Bayesian methods of analysis for cluster randomized trials with binary outcome data. *Stat Med*. 2001;20:453-472.
29. **Spiegelhalter D, Thomas A, Best N.** *WinBUGS 1.3: User Manual*. Cambridge, England: MRC Biostatistics Unit, Institute of Public Health; 2000.
30. **Nunnally J, Bernstein I.** *Psychometric Theory*. New York, NY: McGraw-Hill Inc; 1994.