Title
Variations of physician group profiling indicators for asthma care

Brief title
Variations of physician group profiling

Authors
I-Chan Huang [*], PhD, Gregory B. Diette [‡, §], MD, MHS,
Francesca Dominici [†], PhD, Constantine Frangakis [†], PhD,
Albert W. Wu [*, ‡, §], MD, MPH

Affiliation
[*] Department of Health Policy and Management, Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, Maryland
[†] Department of Biostatistics, Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, Maryland
[‡] Department of Epidemiology, Bloomberg School of Public Health, The Johns Hopkins University, Baltimore, Maryland
[§] Department of Medicine, School of Medicine, The Johns Hopkins University, Baltimore, Maryland

Corresponding author
Albert W Wu, MD, MPH.
Department of Health Policy and Management, Bloomberg School of Public Health
The Johns Hopkins University, Baltimore, Maryland
624 North Broadway
Baltimore, MD 21205-1901
Phone: (410) 955-6567
E-mail: awu@jhsph.edu

Suggested publication category: case study

Prior publication: This manuscript has NOT been published previously (either in whole or in part) NOR be submitted elsewhere in either identical or similar form

Conference Presentation: Presented in part in the 20th Annual Research Meeting of AcademyHealth. June 27-29, 2003. Nashville, Tennessee, USA.

Source of funding: This study was founded by the Pacific Business Group on Health (PBGH)

Word count: 3,638 excluding acknowledge, references, tables and figures

Title

Variations of physician group profiling indicators for asthma care

Abstract

Background

Patient outcomes and consistency of care with guidelines are used as indicators to profile provider performance in asthma care. However, little is know about how much of the variation in profiling measures can be attributed to provider effects, or how reliable are the resulting profiles.

Objectives

To determine how much of the variation in provider profiles can be attributed to provider effects, and how reliable are the indicators for asthma care profiling.

Study design

Cross-sectional study using data from a mailed patient survey. Variations attributable to provider effects are presented using the intraclass correlation coefficient (ICC), estimated using Bayesian hierarchical modeling. The reliability of profiling results was determined using ICC and sample size per physician group.

Participants and settings

Patients with asthma were selected randomly to be surveyed from each of 20 California physician groups between July 1998 and February 1999. A total of 2,515 patients responded.

Main outcome measures

Indicators for physician group profiling included (1) NAEPP guideline-based processes of care, including accessibility of asthma care, self-management knowledge about asthma care, use of inhaled bronchodilators, and use of inhaled steroids, and (2) patient outcomes, including satisfaction with asthma care, improvement in health status, emergency visits, and hospitalizations attributable to asthma.

Results

The variations attributable to provider effects were small (less than 10%) for both process and outcome indicators. For process indicators, self-management knowledge about asthma care had the largest ICC (9.8%), where use of inhaled bronchodilators had the smallest ICC (3.1%). For outcome indicators, satisfaction with asthma care had the largest ICC (9.5%) and hospitalization had the smallest ICC (1.4%). Despite relatively small ICCs, large sample size per physician group yielded acceptable reliability (>0.8) of profiling results, except for use of the inhaled bronchodilators (0.77) and hospitalization (0.60).

Conclusions

The selected indicators for profiling asthma care at the level of physician group were generally reliable, although the ICCs for those indicators were very small. Collecting sufficient case numbers per providers is a key way to achieve the acceptable profiling results.

Key words: asthma; intraclass correlation coefficient; reliability; profiling.

Title

Variations of physician group profiling indicators for asthma care

Introduction

Asthma is a common disease characterized by inflammation of airways and reversible obstruction to airflow. In 1996, an estimated 14.6 million persons in the United States reported having asthma. [1] To bridge the gap between current knowledge and practice and to improve patient outcomes, the National Asthma Education and Prevention Program (NAEPP) Expert Panel, supported by the National Heart, Lung, and Blood Institute (NHLBI), published "Guidelines for the Diagnosis and Management of Asthma" in 1991. [2] The guidelines, which were revised in 1997, [3] emphasize the importance of patient education and appropriateness of medication use.

Currently, many performance and quality oversight organizations (e.g. FACCT, NCQA, PBGH) assess performance of asthma care by physicians, physician groups, and health plans using NAEPP-based guidelines. [4-6] The expectation is that provider profiling can increase accountability of providers to improve quality of care, serve as a tool to control health care costs, and guide consumers to high quality providers. [7] Despite increasing use of profiling to assess provider performance, some important methodological challenges can affect the usefulness of the results.

One of these concerns is how much variation in indicators is across providers. The amount of variation in provider profiles that can be attributed to provider effects after

adjusting for patient case-mix can be estimated using the intraclass correlation coefficient (ICC). [8-11] As the size of ICC increases, quality of care measures are more similar for patients within the same provider, but less similar to those of other providers.

The ICC also reflects the clustering effect of patients nested within providers. Performance indicators with larger ICC suggest that performance measures based on patient observations are not regarded as independent within each provider. The larger ICCs, however, result in sample size calculation based on standard methods not being powerful enough to determine differences in profiling results. To achieve the sufficient power for profiling comparisons, standard sample size estimates need to be inflated by using the inflation factor (IF) (or so-called design effect). [10;12;13] The inflation factor represents the ratio of estimated variances in difference between specialty groups, with and without adjustment for the clustering effect.

Importantly, the ICC and number of patients for each provider (or panel size) together determine the reliability of profiling results. [9;10] Provider profiling will be useful only if provider level variation is important relative to other potential sources of variation (i.e. larger ICC) and reliability is larger. If the variation attributable to provider effects and reliability of profiling results are small, then it is not worthwhile to investing in and disseminating of provider profiles.

To date few studies have examined the sizes of ICCs and reliability of profiling indicators. Based on previous studies, the range of estimated variations attributable to

provider effects varies, depending on type of disease and indicator. In general, the attributable variations to provider effects were very small (usually < 10%). [9;14-17] Hofer et al. who assessed variations of physician profiles for type 2 diabetes found that the overall variance in physician visits and hospitalization rates attributable to differences in physician practice were only 4% and 1%, respectively. [9] Krein and colleagues suggested that the ICCs of process and intermediate outcome indicators for diabetic care ranged from 0-9%. [16] Sixma et al. showed that the ICC of patient satisfaction with GPs was about 5%-10%. [14] A study of Orav and colleagues that assessed ICCs of process of care score demonstrated a wide range of variations due to providers, from 3% (minimum) for cancer screening to 24% (maximum) for management of digoxin.[18] A review by Campbell and others suggested that the ICCs of process indicators were larger than outcomes indicators at the level of individual practice. [19]

Even fewer studies have demonstrated the reliability of profiling indicators. Hofer and colleagues suggested that the reliability of physician profiles for type 2 diabetes for physician visits and hospitalization rates were only 0.41 and 0.17 respectively. [9] A study by Solomon and colleagues that evaluated reliability of performance indicators using the CAHPS survey suggested that the reliability at the level of individual physician ranged from 0.14 to 0.81, and the reliability at level of medical group ranged from 0.15 to 0.81. [17]

In this study, we will address three areas underrepresented in the provider profiling literature. We used consistency with asthma guidelines and patient outcomes as

performance indicators of asthma care. Physician group was used as the unit for profiling. We evaluated (1) how much of the variance of physician group profiling for asthma care is attributed to provider effects? (2) How large is the design effect for physician group profiling? (3) How reliable are physician group profiling of process and outcome indicators? If the variation attributable to provider effects and reliability of profiling results are small, then current profiling practices may need reexamination.

Methods

Study setting

This study was conducted in conjunction with 20 California physician groups that participated in the 1998 Asthma Outcomes Survey (AOS). The AOS was initiated by the Pacific Business Group on Health (PGBH)--a health care purchasing coalition in California, and HealthNet--a California-based health plan, to evaluate, improve and report on the quality of asthma care at physician group level. [20] Experts have suggested the importance of using physician group or medical group as the unit of profiling, instead of health plan. [21] Although health plans may set quality of care policy, most clinical decisions are made by physician groups, which may more directly affect patient outcomes.

Sample selection and data collection

Details on sample selection and data collection have been described. [20] Briefly, the 20 participating physician groups were instructed to use administrative materials to identify all managed care patients with at least one asthma-related encounter in the outpatient, emergency or inpatient settings (identified by ICD-9 code 493.xx) between January 1, 1997 and December 31, 1997. 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 self-administered mailed survey. The instrument was developed largely based on the "Health Survey for Asthma Patients" developed at the Johns Hopkins Health Services Research & Development Center (HSRDC) for the Outcomes Management System (OMS) Consortium Asthma Project of the Managed Health Care Association (MHCA). [22-24] The instrument 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 by PBGH and HealthNet between July 1998 and February 1999 using identical methodology. A total of 2,515 responses were obtained for a response rate of 32.2%.

Performance indicator

Processes of care and patient outcomes were used as asthma care performance indicators for publicly reported physician group comparisons.

Process of care was assessed by consistency of care with the NAEPP asthma guidelines, including accessibility of asthma care, self-management knowledge, use of inhaled steroids, and use of inhaled bronchodilators. Access to asthma care measures accessibility of clinicians or nurse by phone, to make an appointment to see doctors, and to get asthma medications. Self-management knowledge measures ability to manage asthma flares, to appropriately adjust asthma medication, and to correctly identify asthma triggers. For asthma medication use, the NAEPP asthma guidelines advocate inhaled corticosteroids as the most consistently effective long-term control medication for anti-inflammatory. The NAEPP guidelines refer inhaled bronchodilators (or ß-2 agonists) as rescue medications for treatment on an "as needed" basis. [3] Evidence showed that overuse of ß-2 agonists and underuse of inhaled corticosteroids increases the likelihood of hospitalizations, emergency room visits, and death. [25-29] In the survey, patients were rated how many puffs of inhaled bronchodilators and inhaled steroids were used every day. We dichotomized responses for inhaled bronchodilator use into <=8 puffs as "no overuse" and >8 puffs as "overuse," and inhaled steroid use into <=4 puffs as "underuse" and >4 puffs as "no underuse." [23]

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

Risk adjustment

Characteristics of patients and physician groups are potential confounders that may influence physician group performance. However, for profiling, we would like to adjust for the effect of exogenous factors (mainly patient characteristics on which the physician groups have no influence, such as patient's age, sex, education, and baseline severity) rather than endogenous factors (mainly physician group characteristics on which providers can influence, such as physician mix, number of supplementary staff, etc). [30] Adjusting for endogenous factors may mask the true performance of physician groups because these factors can influence the quality of care.

Candidate risk-adjustment variables were collected from the patient survey. Those variables included patient age, sex, education level, type of health insurance, severity, number of asthma-related comorbidity, and health status (the SF-36 physical component

score (PCS) and SF-36 mental component score (MCS). 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 four severity strata (mild-intermittent, mild-persistent, moderate-persistent, and severe-persistent). [3]

Statistical analysis

Chi-square 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 ($p<0.05$) for inclusion in multivariate risk-adjustment models. We include all asthmatics to calculate the ICCs of profiling indicators. However, based on recommendations of the NAEPP guidelines, we only included asthma patients who have moderate-persistent and severe-persistent severity for the inhaled steroid use indicator. [3]

We used Bayesian hierarchical modeling to quantify variations of performance indicators across 20 physician groups that are truly attributed to provider effects. A two-level hierarchical model (level 1 for patients and level 2 for physician groups) was developed to adjust for the clustering effect of patients nested within a specific physician group. The major advantage of hierarchical models is that they allow us to assess provider performance by quantifying random intercepts of logistic regressions at patient level. [31;32] More importantly, hierarchical model can appropriately partition variations of performance measures across physician groups into between-physician group variability

and within-physician group variations by using shrinkage techniques. [11;33] The amount of shrinkage on performance measures for each group is inversely related to variability of observed rate in each group, allowing those groups with small case numbers to shrink their estimates toward the grand mean.

The percentage of variability attributable to provider (groups) effects relative to the overall residual variability can be estimated using the "intraclass correlation coefficient (ICC)." [8-11]

$$\text{ICC} = \frac{\text{Variation between providers}}{\text{Variations between providers and within providers}} \qquad (1)$$

We adopted Turner's method to calculate ICCs for binary performance indicators. [34] The estimations of ICCs under a Bayesian hierarchical approach were carried out using a Markov chains Monte Carlo (MCMC) simulation. The MCMC comprised a burn-in of 500 followed by a further 5000 iterations over which the posterior distribution of ICC were monitored. [35] The results of ICC can be reported as posterior ICC with 95% creditable intervals that allows for direct estimation of the probability of ICC. [36]

We calculated the inflation factor (IF) to reflect the desired sample size. [10;12;13] The inflation factor is:

$$IF = 1+ (n-1) * ICC \quad (2)$$

Where the n is average sample size of each physician group. If the ICC is zero, suggesting no clustering effects of patient nested in physician group, then IF is zero and no inflation of sample size is needed.

We calculated reliability by combining the information of ICC and sample size of each physician group using the following formula. [9;10]

$$\text{Reliability} = \frac{n * ICC}{1 + (n-1) * ICC} \quad (3)$$

Based on formula (3), we can further calculate the required sample size based on expected reliability of profiling results. To date, there is still no acceptable standard for judging reliability of provider profiling. Most of studies suggest that the reliability should be larger than 0.8. [9;16] Another suggests the acceptable reliability at the level of 0.7.[37]

Statistical package

STATA 7 was used for bivariate analyses and Winbugs 1.3 for ICC calculation. [35]

Results

Characteristics of physician groups and respondents

Of the 20 participating physician groups, 8 were located in Northern California and 12 in Southern California. The numbers of case in each physician groups ranged from 63 to 218 [mean: 125.8]. Characteristics of the 2,515 asthma patients who participated in this study are shown in Table 1. Patients ranged in age from 18-56 years with a mean of 39.9 years [SD: 9.5]. 71.2% were female; 70.3% were white, and 5.1% were African-American; 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 number of comorbidities was 2.1 [SD: 1.4].

Variations in performance indicators attributable to provider effects across physician groups

In general, ICCs using Bayesian hierarchical modeling indicated that for both process and outcome indicators, variations attributable to provider effects were small (less than 10%) (Table 2). Indicators of guideline consistency demonstrated slightly larger ICCs than outcome indicators. Among indicators of guideline consistency, self-management knowledge about asthma care had the largest ICC (9.83%), while use of inhaled bronchodilators had the smallest ICC (3.08%). Among patient outcome indicators,

satisfaction with asthma care had the largest ICC (9.53%) and hospitalization had smallest ICC (1.35%).

Inflation factor calculation

Figure 1 demonstrates that there were positive linear relationships between physician group sizes and IFs. The magnitudes of the linear relationships, however, depended on the ICCs. Given the same group size, those performance indicators with larger ICC usually had larger IFs. For example, given the 126 patients per physician, for indicators of guideline consistency, self-management knowledge about asthma care had the largest IF (13.3), and the use of inhaled bronchodilators had the smallest IF (4.9). For patient outcome indicators, satisfaction with asthma care had the largest IF (13.0), and hospitalization had smallest IF (2.7). The discrepancies of IFs among performance indicators increased as the group sizes increased.

Reliability of profiling results

Table 2 shows the reliability of profiling results. In general, the reliability of profiling for asthma care at the level of physician group was acceptable based on the criteria of >= 0.8. Indicators of consistency of care with guidelines demonstrated slightly larger reliability than outcome indicators. The range of reliability is from 0.60 to 0.92. Six out of selected profiling indicators had the reliability >=0.80, except inhaled bronchodilator use and hospitalization. Self-management knowledge about asthma care had the largest reliability (0.92), and patient satisfaction with asthma care had the second largest

reliability (0.91). In contrast, inhaled bronchodilator use and hospitalization were less reliable, 0.77 and 0.60 respectively.

Sample size needed to achieve reliable profiling

Figure 2 demonstrates the relationship between reliability of profiling and desired sample size. Given a fixed ICC, the relationship between reliability of profiling and sample size per physician group was exponential. Generally speaking, indicators that had smaller ICCs usually needed a larger sample size per group (compared to current sample size per group of 126) to achieve the acceptable reliability. If we set the reliability at the level of 0.7 and 0.8, greater sample size (170 and 292 per group respectively) was needed only for the hospitalization indicator. If we assumed a strict reliability level at 0.9, greater sample size was needed for indicators of accessibility, bronchodilator uses, emergency department visit, and hospitalization visit (159, 283, 216, and 658 per group respectively).

Discussion

Quality of care and performance oversight organizations are starting to use consistency of care with asthma guidelines and patient outcomes as indicators of provider performance for asthma care. However, the usefulness of these indicators may be limited by the proportion of variation in these performance indicators that is attributed to provider

effects. In addition, we do not have much information regarding the reliability of profiling results. This study provides evidence to address these issues.

Our results showed that, in general, variations in selected performance indicators attributable to provider effects (ICCs) in asthma care across 20 physician groups were small. Among these indicators, self-management of asthma care and satisfaction with asthma care had larger attributable performance variations. In contrast, hospitalization, use of inhaled bronchodilators, and emergency room visits had less variation attributable to provider effects. The larger ICCs for the former indicators may in part reflect homogenous asthma management styles within physician groups, and more heterogeneity across physician groups. However, the smaller ICCs of the later indicators may reflect the characteristics of asthma treatment, which can be significantly affected by patient characteristics (e.g. asthma severity) and is indirectly related to management of physician groups.

In this study the variation attributable to provider effects for inhaled bronchodilator use was smaller (less than 60%) than for inhaled steroid use. This might suggest that although inhaled steroids have been strongly recommended by the NAEPP guidelines as being the most consistently effective long-term control medication for many years, the guideline consistency with inhaled steroid use across physician groups is still worse than for inhaled bronchodilator uses. [2;3;35]

The size of variation attributable to provider effects is an important factor that can affect usefulness of report cards. In this study, despite using sophisticated methods to quantify variations attributable to provider effects, the resulting smaller ICCs in part may reflect relatively large random variation, suggesting apparent variations in provider performance when there are no more than would be expected by chance alone (i.e. small signal-to-noise phenomena). [9;18;38]

Another use of ICC calculation is to determine the inflation effect (IF). When we compare profiling results across providers, the clustering effect of patients nested within physician groups usually causes a loss of statistical power. [10;12;13] The IF can provide guidance for how much additional sample size is needed to achieve the equivalent statistical power of standard sample size calculation. The IF of 1 implies no group-level clustering effect, and IF of > 1 implies a clustering effect. Based on our findings, indicators with higher group-level of clustering, such as self-management knowledge, need to require larger sample sizes to compensate the loss of statistical power.

In this study, we also demonstrated the reliability of profiling indicators for asthma care at the level of physician group. We found that the reliability of these profiling indicators was generally good (>0.8), with the exception of inhaled bronchodilator use and hospitalization. This finding contrasts with previous studies that demonstrated low reliability of profiling results at the level of individual physician. [9;16]

There are some implications from our findings. First, the ICC calculation is the key element of provider profiling, as it helps to determine the reliability of provider profiling and the design effect. Without precise information in the ICC, we cannot decide how reliable the provider profiling is, and what sample sizes are needed when comparing profiling results. Quality of care and performance oversight organizations should calculate the ICCs for different profiling indicators and for different levels of profiling.

Second, the FACCT, NCQA, PBGH and the other organizations that engage in profiling providers for asthma care need to be careful in selecting performance indicators. In particular, they need to select profiling indicators based on larger ICC and greater reliability. For process indicators of asthma care, our findings suggest that knowledge about asthma care and use of inhaled steroids may be good performance indicators for physician group profiling. For outcome indicators, satisfaction with asthma care and improvement in health status may be appropriate. It seems less advisable to use inhaled bronchodilators and hospitalization as indicators.

Third, a sufficient sample size appears to be the key solution to achieve acceptable reliability of profiling results. However, increasing sample size for each provider may be not efficient or possible. It is particularly difficult for profiling individual physicians for specific diseases with lower prevalence rate. Hofer and colleagues demonstrated that to achieve reliability of 0.8 for profiling hospitalization and physician visit rates of diabetic care, we need at least 100 patients for each provider, [9] suggesting that profiling at the

level of individual physician might be not feasible. Based on our study, physician group may be a more feasible unit of profiling than individual physician.

In interpreting our findings, several potential limitations should be noted. First, there was a low response rate to the patient survey. It is possible that the study sample may be biased to include patients who had better outcomes, and may not be representative of all the patients in the population. Second, the variation attributable to provider effects may be underestimated because of unmeasured confounders not included in our models. Furthermore, we did not collect clinical assessments or non-patient characteristics, such as supply of providers or hospitals in market. Lack of adjustment for these factors may potentially increase random variation. [10;11] On the other hand, the estimated variation attributable to provider effects may be overestimated because we cannot precisely partition overall variations into physician group level. Based on clustering characteristics between patients, physicians, and physician groups, it is better to partition overall variations into three levels. However, our data do not allow us to further partition variations into physician and physician group.

In conclusion, for performance profiling across 20 physician groups, the variations attributable to provider effects were small. However, the reliability of profiling results was generally acceptable because of sufficient case numbers for each physician group. For profiling we recommend use of indicators with larger reliability.

REFERENCES

1. US DHHS: Action Against Asthma. Washington, DC: US DHHS; 2000.

2. National Heart Lung and Blood Institute: National Asthma Education Program: Guidelines for the Diagnosis and Management of Asthma. Bethesda, Md: US DHHS; 1991.

3. National Heart Lung and Blood Institute: Guidelines for the Diagnosis and Management of Asthma: Expert Panel Report No. 2. Bethesda, Md: US DHHS; 1997.

4. Pacific Group on Health. Available at http://www.pbgh.org . Accessed August 8, 2003.
5. Foundation for Accountability. Available at http://www.facct.org/facct/site/facct/facct/home . Accessed August 8, 2003.
6. National Committee for Quality Assurance: The State of Health Care Quality: 2002. Washington, DC: National Committee for Quality Assurance; 2002.

7. 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.

8. Gulliford MC, Ukoumunne OC, Chinn S. Components of variance and intraclass 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.

9. 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.

10. Snijders TAB, Bosker RJ. Multilevel analysis an introduction to basic and advanced multilevel modeling. London: Sage Publications; 1999.

11. Raudenbush SW, Bryk AS. Hierarchical linear models: applications and data analysis methods. Thousand Oaks, CA: Sage publications; 2002.

12. Campbell MK, Mollison J, Grimshaw JM. Cluster trials in implementation research: estimation of intracluster correlation coefficients and sample size. Stat Med 2001:20:391-399.

13. 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.

14. Sixma HJ, Spreeuwenberg PM, van der Pasch MA. Patient satisfaction with the general practitioner: a two-level analysis. Med Care 1998:36:212-229.

15. 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

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. Orav EJ, Wright EA, Palmer RH, Hargraves JL. Issues of variability and bias affecting multisite measurement of quality of care. Med Care 1996:34:SS87-101.

19. Campbell M, Grimshaw J, Steen N. Sample size calculations for cluster randomised trials. J Health Serv Res Policy 2000:5:12-16.

20. Masland M., Wu AW, Diette GB, Dominici F, Skinner EA, Esquivel M. The 1998 Asthma Outcomes Survey. San Francisco, CA: Pacific Business Group on Health; 2000.

21. Epstein AM. Rolling down the runway: the challenges ahead for quality report cards. JAMA 1998:279:1691-1696.

22. Steinwachs DM, Wu AW, Skinner EA. How will outcomes management work? Health Aff (Millwood) 1994:13:153-162.

23. Diette GB, Wu AW, Skinner EA, Markson L, Clark RD, McDonald RC, Healy JP, Jr., Huber M, Steinwachs DM. Treatment patterns among adult patients with

asthma: factors associated with overuse of inhaled beta-agonists and underuse of inhaled corticosteroids. Arch Intern Med 1999:159:2697-2704.

24. Wu AW, Young Y, Skinner EA, Diette GB, Huber M, Peres A, Steinwachs D. 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. Donahue JG, Weiss ST, Livingston JM, Goetsch MA, Greineder DK, Platt R. Inhaled steroids and the risk of hospitalization for asthma. JAMA 1997:277:887-891.

26. Blais L, Ernst P, Boivin JF, Suissa S. Inhaled corticosteroids and the prevention of readmission to hospital for asthma. Am J Respir Crit Care Med 1998:158:126-132.

27. Nestor A, Calhoun AC, Dickson M, Kalik CA. Cross-sectional analysis of the relationship between national guideline recommended asthma drug therapy and emergency/hospital use within a managed care population. Ann Allergy Asthma Immunol 1998:81:327-330.

28. Suissa S, Ernst P, Benayoun S, Baltzan M, Cai B. Low-dose inhaled corticosteroids and the prevention of death from asthma. N Engl J Med 2000:343:332-336.

29. Adams RJ, Fuhlbrigge A, Finkelstein JA, Lozano P, Livingston JM, Weiss KB, Weiss ST. Impact of inhaled antiinflammatory therapy on hospitalization and

emergency department visits for children with asthma. Pediatrics 2001:107:706-711.

30. Welch HG, Black WC, Fisher ES. Case-mix adjustment: making bad apples look good. JAMA 1995:273:772-773.

31. 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.

32. Katon W, Rutter CM, Lin E, Simon G, Von Korff M, Bush T, Walker E, Ludman E. Are there detectable differences in quality of care or outcome of depression across primary care providers? Med Care 2000:38:552-561.

33. Cowen ME, Strawderman RL. Quantifying the physician contribution to managed care pharmacy expenses: a random effects approach. Med Care 2002:40:650-661.

34. Turner RM, Omar RZ, Thompson SG. Bayesian methods of analysis for cluster randomized trials with binary outcome data. Stat Med 2001:20:453-472.

35. Spiegelhhalter D, Thomas A, Best N. Winbugs 1.3: User Mannual. Cambridge, UK: MRC Biostatistics Unit, Institute of Public Health; 2000.

36. Gelman A, Carlin JB, Stern HS, Rubin DB: Bayesian Data Analysis. Roca Raton, FL: Chapman & Hall/CRC; 2000.

37. Nunnally J, Bernstein I. Psychometric Theory. New York, NY: McGraw-Hill, Inc.; 1994.

38. Shwartz M, Ash AS, Anderson J, Iezzoni LI, Payne SM, Restuccia JD. Small area variations in hospitalization rates: how much you see depends on how you look. Med Care 1994:32:189-201.

Table 1: Characteristics of patients with asthma (n=2,515)

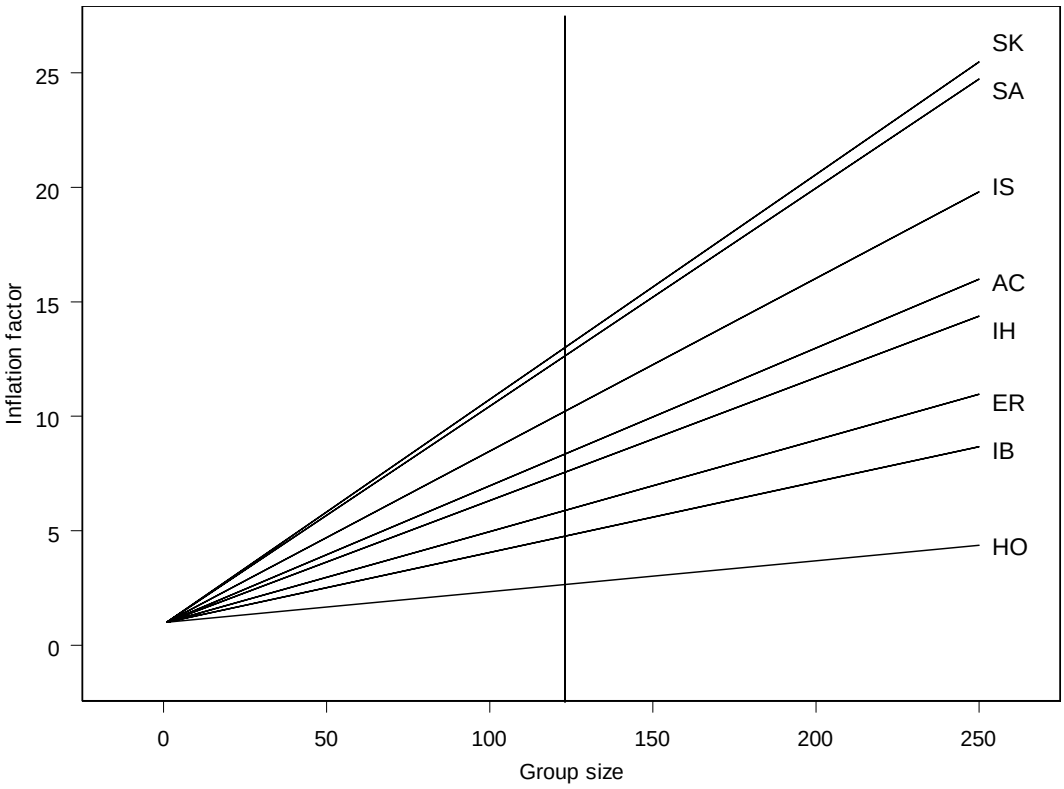
	Percentage or mean (SD)
Age, %	
18-24	7.2
25-34	22.0
35-44	34.6
45-54	33.2
55 and above	3.1
Overall: mean (SD)	39.9 (9.5)
Sex, %	
Females	71.2
Education, %	
High school or below	18.4
College	65.3
Graduate	16.3
Health Insurance status, %	
Private---through employer	69.1
Private---through self-purchase	24.8
Public---Medicare, Medicaid	1.3
Others	4.9
Drug insurance coverage, %	96.5
Asthma severity, %	
Mild intermittent	14.4
Mild persistent	19.2
Moderate persistent	49.3
Severe persistent	17.1
Number of comorbidity, mean (SD)	2.1 (1.4)
Selected comorbidity	
Sinuitis	38.0
Heart burn	31.2
Bronchitis	14.3
Health status–SF36 component scores	
Physical component score (PCS), mean (SD)	45.7 (10.3)
Mental component score (MCS), mean (SD)	47.4 (10.7)

Table 2: Variations in performance indicators attributable to differences among group practice

	ICC %	95% C.I.	Reliability
Guideline consistency			
Any access barrier to asthma care	5.37	3.34- 8.50	0.85
Knowledge about asthma care	9.83	6.47-14.64	0.92
Bronchodilator inhaler	3.08	1.71- 5.22	0.77
Steroid inhaler ¶	7.55	4.35-12.08	0.89
Patient outcomes			
Satisfaction with asthma care	9.53	6.14-14.28	0.91
Improved health status	6.02	3.75- 9.33	0.87
Emergency visit due to asthma	4.00	2.38-6.28	0.81
Hospitalization due to asthma	1.35	0.65- 2.37	0.60

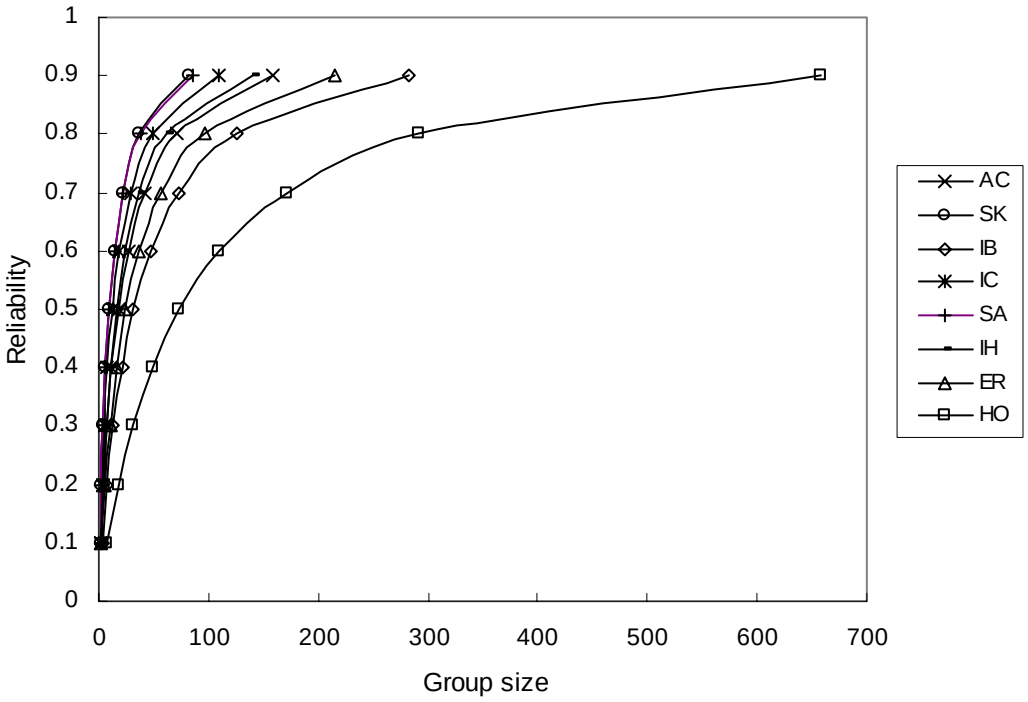
¶ Asthmatics with moderate-persistent or severe-persistent severity

Figure 1: The effect of ICC and panel size on inflation factor



The vertical line indicated the average group size (126) in this study.
AC: Access to asthma care (ICC: 0.054). SK: Self-management knowledge (0.098). IB: Inhaled bronchodilators (0.031). IS: Inhaled steroids (0.076). SA: Satisfaction with asthma care (0.095). IH: Improvement in health status (0.060). ER: Emergency room visit (0.04). HO: Hospitalization (0.013).

Figure 2: Relationship between sample size of each physician group and reliability



AC: Access to asthma care (ICC: 0.054). SK: Self-management knowledge (0.098). IB: Inhaled bronchodilators (0.031). IS: Inhaled steroids (0.076). SA: Satisfaction with asthma care (0.095). IH: Improvement in health status (0.060). ER: Emergency room visit (0.04). HO: Hospitalization (0.013).