Predicting Persistently High Primary Care Use

James M. Naessens, MPH
Macaran A. Baird, MD, MS
Holly K. Van Houten, BA
David J. Vanness, PhD
Claudia R. Campbell, PhD

1Department of Health Sciences Research, Division of Health Care Policy & Research, Mayo Clinic College of Medicine, Rochester, Minn
2Department of Family Medicine and Community Health, University of Minnesota, Minneapolis, Minn
3Department of Population Health Sciences, University of Wisconsin, Madison, Wisc
4Department of Health Systems Management, School of Public Health and Tropical Medicine, Tulane University, New Orleans, La

ABSTRACT

PURPOSE We wanted to identify risk factors for persistently high use of primary care.

METHODS We analyzed outpatient office visits to practitioners in family medicine, general internal medicine, general pediatrics, and obstetrics for 1997-1999 among patients in a small Midwestern city covered by a fee-for-service insurance plan with no co-payments for physician visits and no requirement for referral to specialty care. Logistic regression was used to predict which patients with 10 or more primary care visits in 1997 would repeat high use in 1998 based on demographic and diagnostic categories (adjusted clinical groups [ACGs]). A confirmatory data set (high primary care use in 1998 persistent into 1999) was used to evaluate the model.

RESULTS Two percent of the 54,074 patients had 10 or more primary care visits in 1997, and of these, almost 19% had 10 or more visits in the next year. Among adults, 4 ambulatory diagnosis groups (ADGs) were simultaneously positive predictors of repeated high primary care visits: unstable chronic medical conditions, see and reassure conditions, minor time-limited psychosocial conditions, and minor signs and symptoms. Meanwhile, pregnancy was negatively associated. The area under the receiver operating characteristic (ROC) curve was 0.794 for adults in the developmental data set and 0.752 in the confirmatory data set, indicating a moderately accurate assessment. A satisfactory model was not developed for pediatric patients.

CONCLUSIONS Many persistently high primary care users appear to be overserved but underserved, with underlying problems not addressed by a medical approach. Some may benefit from psychosocial support, whereas others may be good candidates for disease management interventions.

Ann Fam Med 2005;3:324-330. DOI: 10.1370/afm.352.

INTRODUCTION

The well-known concentration of health care utilization and costs among relatively few individuals allows for targeted interventions and modeling of patient risks for more equitable health care reimbursement. Factors associated with high utilization include patient demographics, previous use of health care, patient diagnosis, and severity of illness. If certain patient characteristics are predictive of high persistent use, it may be possible to offer more cost-effective alternatives to frequent primary care visits, including disease management, case management, group visits, and patient education. Better management of persistent primary care use may better address the patients' underlying problems, reduce unnecessary demand, and relieve some of the pressure on the capacity of primary care providers to deliver care to all patients.

Most efforts to identify and adjust for patients with high expected health care use have been directed at predicting total health care expenditures or those associated with specific chronic diseases. This study focuses on characteristics of patients with high primary care use. Primary care provides integrated, accessible health care services by cli-
nicians who are responsible for a great majority of personal health care needs, develop a sustained partnership with patients, and provide care in collaboration with the family and the community. Patients seek health services based on perceived needs, self-assessed health status, and expected benefits of care. Moreover, primary care often serves as an access point to more costly medical care services. Excessive demand for and unlimited use of primary care services, if unnecessary or ineffective, may constrain primary care capacity.

We characterize high and persistent primary care users as overserviced but underserved, because their needs are not being met through increased medical care use. Some evidence suggests that a large proportion of such patients seek help on nonmedical issues or for minor health concerns. Such patients may be amenable to interventions to reduce ineffective or unnecessary use of primary care.

**METHODS**

**Study Site and Patient Population**

The study population included approximately 58,000 people continuously insured with a fee-for-service plan in a small urban area in the midwestern United States. Patients who did not authorize research use of their medical records were excluded (6% of eligible members), resulting in a study sample of 54,074 patients. All outpatient office visits (CPT-4 codes: 99201-99499) were selected for 1997-1999 from billing data. These visits were subdivided into primary care visits and specialty care visits. Visits to practitioners in family medicine, general internal medicine, general pediatrics, and obstetrics were considered primary care visits. We defined high primary care use as 10 or more visits in the calendar year and persistently high use as having 10 or more annual visits to primary care in 2 consecutive years. Insurance coverage provided first dollar coverage for all physician visits (ie, no co-payments), and no referrals were required to receive specialty care. Available demographic data included patient age, sex, and employee or dependent status. Nine hundred twenty-nine high primary care users in 1997 who were eligible in 1998 were used to develop the model to predict high persistent use. Only 58 patients with high use in 1997 who were eligible for services in both 1997 and 1998. Candidate independent variables included the 34 ambulatory diagnosis groups (ADGs) instead of the terminal classification (ACGs). ADGs are created by combining individual ICD-9-CM diagnosis codes into clusters based on several factors including: need for specialty care, chronicity, and severity.

The ICD-9-CM diagnosis codes were captured from administrative billing data for all services for patients in the study. Each outpatient visit contained up to 3 diagnosis codes; hospital services were summarized into encounters including up to 16 diagnosis codes. We assembled the list of all unique diagnosis codes over all items of utilization for each patient for each calendar year. ACG software (version 4.1) was applied to collapse diagnoses for each patient into the 34 ADG indicators.

**Statistical Methods**

A multivariate logistical regression model was developed to identify risk factors associated with persistently high primary care use among those patients with high primary care use in 1997 who were eligible for services in both 1997 and 1998. Candidate independent variables included the 34 ADGs, number of ADGs, age categories, employee or dependent status, and sex. Pediatric patients (less than 17 years old) were analyzed separately. Based on the estimated odds ratio for risk factors that were significant in the final model using the developmental data set (1997 predicting 1998 use), we constructed a scoring algorithm to be tested on the confirmatory data set (1998 predicting 1999 use). A goodness-of-fit test of this model was performed. Sensitivity and specificity statistics (as well as receiver operating characteristic [ROC] curve analysis) for predicting persistently high use based on the scoring algorithm were assessed for both the developmental and confirmatory data sets. Furthermore, the confirmatory results were stratified by patients who had 3 years of high primary care use are different from those whose high use only persists for 2 years.

**Clinical Risk Factors**

Several models for predicting future health service costs and associated charges have been used for risk-adjusting capitation and payment rates. For this study, we used the Johns Hopkins adjusted clinical groups (ACGs) system to categorize patients into diagnosis groups. ACGs were developed by a team including medical experts in primary care to be both clinically meaningful and predictive of utilization and resource costs. ACGs are currently used by managed care plans and others for risk-adjusting financial arrangements with providers and for physician profiling. Because the focus of this study is on predicting persistently high-use patients of primary care visits rather than high-use patients of all types of ambulatory care services, we used the 34 ambulatory diagnosis groups (ADGs) instead of the terminal classification (ACGs). ADGs are created by combining individual ICD-9-CM diagnosis codes into clusters based on several factors including: need for specialty care, chronicity, and severity.

The ICD-9-CM diagnosis codes were captured from administrative billing data for all services for patients in the study. Each outpatient visit contained up to 3 diagnosis codes; hospital services were summarized into encounters including up to 16 diagnosis codes. We assembled the list of all unique diagnosis codes over all items of utilization for each patient for each calendar year. ACG software (version 4.1) was applied to collapse diagnoses for each patient into the 34 ADG indicators.
RESULTS

Two percent (n = 987) of the study population of 54,074 patients had 10 or more primary care visits in 1997. This small group of pediatric and adult patients accounted for 18% of all primary care visits (Figure 1), as well as 11% of the total paid insurance claims in that year. Among the 15,175 pediatric patients seen in 1997 and eligible in 1998, 134 (0.9%) accounted for 7.7% of all visits and used 12.6 visits per patient compared with 1.4 visits per patient among all pediatric patients. Among adult patients, 795 (2.6%) of 31,034 accounted for 20.7% of all visits or 13.5 visits per patient compared with 1.7 visits per patient overall. High-visit patients’ differences on several demographic and visit characteristics are displayed in Table 1. The most frequent diagnoses identified at the primary care visit for these 2 groups are displayed in Supplemental Appendix 1, available online-only at http://www.annfammed.org/cgi/content/full/3/4/324/DC1.

Among the 987 high-use patients in 1997, 929 (94.1%) were also eligible for health care coverage in 1998, compared with 85.3% continued eligibility among the rest of the 1997 eligible patients. The reason for loss of insurance coverage for these 58 patients is not known, but the termination rate was
not higher than among non–high-visit primary care patients. In 1997 we found 23.9% (32 of 134) of high-use primary care pediatric patients were also high-use patients in 1998. Likewise 141 of the 795 (17.7%) adult high-use patients continued to have high primary care use in 1998. Younger children and older adults appear to be more likely to have persistently high use (Supplemental Appendix 2, available online-only at http://www.annfammed.org/cgi/content/full/3/4/324/DC1). No difference for persistently high use by sex of the patient was observed among pediatric patients (25.9% among females and 22.5% among males), but male adults were more likely to be persistently high-use patients than female adults (32 of 75, 42.7% males; 109 of 720, 15.1% females).

Table 2 provides the odds ratios and 95% confidence intervals of the logistic regression model estimated for adults in both 1997 and 1998 for significant risk factors associated with high primary care visit use in both years. The model identified 4 ADGs that significantly increased the risk of a patient being a persistently high user of primary care visits. These ADGs were chronic medical-unstable (11); see and reassure (30); psychosocial: time limited, minor (23); and signs/symptoms, minor (26). ADG 33 for pregnancy was negatively associated with the persistence of clinic visits. Once these medical conditions were considered, age-group and sex were not statistically significant at $P < .05$. Table 2 also reports the score assigned to each risk factor based on the sign and magnitude of the estimated odds ratio. Both chronic medical-unstable (ADG 11) and see and reassure (ADG 33) conditions more than doubled the odds of high persistent visit use and were assigned a score of 2, whereas psychosocial (ADG 23) and signs/symptoms (ADG 26) increased the odds of high use in both years by 50% and were given a score of 1. Pregnant patients were assigned a score of -4. Patients who were high-use patients in 1997 but who did not have any of these risk factors were assigned a score of 0 in the model.

We tested the ability of the scoring model to predict high usage in 2 ways. First, we examined how well the risk factors predicted which high-use patients in 1997 would continue to be high-use patients in 1998. Next, we tested the model on a confirmatory sample to determine whether the model applied beyond the original patient sample.

Figure 2 provides the percentage of high-use patients in 1997 that

![Figure 2. Persistent high primary care use by model score for adult and pediatric patients in 1998.](image-url)
the scoring model predicted correctly to be high-use patients in 1998. Patients with a score of 1 or greater were correctly identified for future high use with a sensitivity of 80.3% and a specificity of 62.7%. In other words, 80.3% of patients with a score of 1 or greater in 1997 had 10 or more primary care visits in 1998, whereas 62.7% of patients with scores less than 1 did not have 10 or more visits in 1998. If the threshold score were set at 2 or greater, the sensitivity was lowered to 50.3% and the specificity was increased to 81.2%. The area under the ROC curve was 0.794 for adults, indicating a moderately accurate assessment. Furthermore, the Hosmer and Lemeshow goodness-of-fit test was not statistically significant \((P = .0992)\) implying that the model's estimates fit the data at an acceptable level. The scoring model was much less accurate in predicting persistence among pediatric patients.

The same general pattern was observed on 1998 data to predict 1999 high use (Figure 3). Among adults, 873 who were high-use patients in 1998 were enrolled in 1999. Of these, 132 were also high-use patients in 1997. Among pediatric patients, 237 high-use patients in 1998 were enrolled in 1999, including 31 children who were high-use patients in 1997. Of the 237 high-use pediatric patients in 1998, 25.3% \((n = 60)\) also had high use in 1999. Among high-use adults, 19.6% \((171 \text{ of } 873)\) were persistently high-use patients across both years. With a score of 1 or greater, the sensitivity of the model to detect persistently high-use patients was 75.8% and the specificity was 57.9%. A score of 2 or greater resulted in sensitivity of 49.8% and a specificity of 80.0%. The area under the ROC curve for adults in the confirmatory data set was 0.752 \(0.730\) for the 741 new high-use patients and 0.580 for the 132 continuing high-use patients).

Interestingly, the model was less successful in predicting future high use for 163 patients who were persistently high-use patients in 1997 and 1998. The lower prediction suggests that different factors may be at play among these 2 groups.

**DISCUSSION**

Not surprisingly, our study showed a small proportion of primary care patients (2%) consumed a large share of total primary care visits (18%). Additionally, we found that a considerable portion of these patients persisted in their relatively high use for 2 years. A multivariate logistic model to predict persistently high use found patients with unstable chronic medical conditions (eg, chronic obstructive lung disease and rheumatoid arthritis), time-limited minor psychosocial conditions (eg, other acute reactions to stress [ICD-9 diagnosis code 308.3]), minor signs and symptoms (eg, headache), and see and reassure conditions, including breast hypertrophy, were more likely to have 10 or more primary care visits in 2 consecutive years. Furthermore, in the 1998-1999 confirmatory analysis, the scoring model consistently identified a sizable portion of the persistently high-use patients among new high-use patients and continuing high-use patients.
Correct prediction of which primary care patients are likely to continue to demand high levels of primary care enables health plans and clinicians to focus on interventions that not only are cost-effective but also may improve both patient and clinician satisfaction with the care process as well as overall health outcomes. Patients who have unstable chronic medical conditions have been shown to be good candidates for disease management efforts. As has been found in Sweden and the United States, a large proportion (58%) of patients with high primary care use in our study have diagnoses suggestive of psychophysologic conditions. Because these patients have many visits but keep returning for relatively minor medical conditions, we label them as “overserviced, underserved.” Peltenburg et al found that every sixth (15.8%) consultation revealed emerging psychosocial agenda.

The existence of persistent-use patients who consume a large proportion (8.7%) of primary care capacity and have relatively minor conditions also has clinical or managed care relevance. We speculate that many of these patients may be undergoing stressful conditions with little social support. They may need psychological or spiritual counseling, group therapy, or more social activity and involvement. Rather than addressing their true needs or the underlying cause of distress, the medical care system may try to alleviate the patient’s physical symptoms. When these palliative measures prove unfruitful, the patients are often labeled as uncooperative, noncompliant, or psychosomatic. This situation not only leads to unhappy patients with unresolved issues, but creates dissatisfaction and frustration among the physicians who are trying to address patient needs as well as they can. If these overserviced, underserved patients are indeed suffering from nonmedical problems that can be identified, alternative social support services or integrated consultations with primary care physicians can be offered to the patient and close family members to better address patient needs through nonmedical approaches. Currently, physicians serving this subgroup are caught in a series of unsatisfying interactions in which more diagnostic testing or medical referrals seem to be the only options. Furthermore, access and time resources for other patients with more straightforward medical dilemmas and illnesses are being consumed, leaving less time and energy for prevention, chronic illness care, and coordination of medically complex care. This opportunity cost is considerable but difficult to calculate.

Do 10 or more primary care visits indicate inappropriate use or ineffective care? Assignment of any threshold certainly can be considered arbitrary. We have shown that, in this population, only 2% of patients had a level of primary care visits this high. They accounted for more than 18% of all primary care visits. Similar results for thresholds of 8, 12, and 15 visits were 3.4%, 1.2%, and 0.5% of patients with 24%, 12%, and 6% of visits, respectively. For the purpose of determining which patients might merit special interventions, we focused our attention as narrowly as possible while still including a substantial proportion of primary care visits. The finding that these patients share common diagnostic characteristics leads to the possibility that interventions can be found that are beneficial and cost effective.

Our study is limited by a single patient group in the Midwest with fee-for-service medical care. During the study period, all patients had medical insurance with no co-payments or coinsurance to limit use. Health care use by overserviced, underserved patients may be less in managed care or other settings where primary care use requires a co-payment. Furthermore, our identification of predictors for persistently high primary care use was limited to basic demographic factors and the categories available in the ACG system, which relies on ICD-9 codes from billing data with their inherent problems. It is worthwhile to note that we were able to show the model was consistent for an independent sample.

Further research is needed to test our hypothesis that patients with persistently high use in 1 year whose diagnoses fall within these ADGs indeed suffer from nonmedical conditions that are amendable to alternative interventions. Other information that might be helpful in determining reasons for persistently high use among these adult patients includes screening for depression and subthreshold somatoform disorders. For children, information about their families and parents would likely be more useful. Finally, future research should tell us whether these proposed efforts at reducing primary care use among the persistently high-use patients are cost-effective.

A small proportion of patients consume a large proportion of primary care visits. We developed a predictive model using ADGs that was successful in predicting which high-use patients persisted in high primary care use the next year. Among these persistently high-use primary care patients were a substantial number of patients with unstable chronic medical conditions who are candidates for disease management efforts and a large proportion of patients with psychophysologic diagnoses who may benefit from more directed, formal, and organized approaches to treat psychological problems, reduce stress, and address nonmedical issues.

To read or post commentaries in response to this article, see it online at http://www.annfammed.org/cgi/content/full/3/4/324.

Key words: Ambulatory care diagnosis groups; forecasting; primary health care/utilization; clinic visits
Submitted October 5, 2004; submitted, revised, March 9, 2005; accepted March 16, 2005.

Acknowledgment: The authors thank Marlene Boyd for editorial support and graphic design in the preparation of this manuscript.

References

1. Health Care Financing Administration. Medicare: a profile. Washington, DC: United States Department of Health & Human Services; 1995.
2. Riley G, Tudor C, Chiang YP, Ingber M. Health status of Medicare enrollees in HMOs and fee-for-service in 1994. Health Care Financ Rev. 1996;17:65-76.
3. Iezzoni LI, Ayanian JZ, Bates DW, Burstin HR. Paying more fairly for Medicare capitated care. N Engl J Med. 1998;339:1933-1938.
4. Kuttner R. The risk-adjustment debate. N Engl J Med. 1998;339:1952-1956.
5. Iezzoni LI. Risk and outcomes. In: Risk Adjustment for Measuring Healthcare Outcomes. 2nd ed. Iezzoni LI, ed. Chicago, Ill: Health Administration Press; 1997.
6. Reuben DB, Keeler E, Seeman TE, et al. Development of a method to identify seniors at high risk for high hospital utilization. Med Care. 2002;40:782-793.
7. Riegel B, Carlson B, Kopp Z, et al. Effect of a standardized nurse case-management telephone intervention on resource use in patients with chronic heart failure. Arch Intern Med. 2002;162:705-712.
8. Roggenkamp SD, White KR. Is hospital case management a rationalized myth? Soc Sci Med. 2001;53:1057-1066.
9. Schore JL, Brown RS, Cheh VA. Case management for high-cost Medicare beneficiaries. Health Care Financ Rev. 1999;20:87-101.
10. Goldstein R. The disease management approach to cost containment. Nurs Case Manag. 1998;3:99-103.
11. Masley S, Sokoloff J, Hawes C. Planning group visits for high-risk patients. Fam Pract Manag. 2000;7:33-37.
12. Friedhoff SG. Intensive case management of high-risk patients in a family medicine residency setting. J Am Board Fam Pract. 1999;12:264-269.
13. Lynch JP, Forman SA, Graff S, Gunby MC. High-risk population health management–achieving improved patient outcomes and near-term financial results. J Manag Care. 2000;6:781-791.
14. Ash A, Porell F, Gruenberg L, Sawitz E, Beiser A. Adjusting Medicare capitation payments using prior hospitalization data. Health Care Financ Rev. 1989;10:17-29.
15. Meenan RT, O’Keeffe-Rosetti C, Hornbrook MC, et al. The sensitivity and specificity of forecasting high-cost users of medical care. Med Care. 1999;37:815-823.
16. Verbrugge LM, Patrick DL. Seven chronic conditions: their impact on US adults’ activity levels and use of medical services. Am J Public Health. 1995;85:173-182.
17. Hoffman C, Rice D, Sung HY. Persons with chronic conditions. Their prevalence and costs. JAMA. 1996;276:1473-1479.
18. Donaldson MS, Yordy KD, Lohr KN, Varselov NA, eds. Primary Care: America’s Health in a New Era. ed. Washington, DC: National Academies Press; 1996.
19. Lucas SF, Peek CJ. A primary care physician’s experience with integrated behavioral health care: what difference has it made? In: Behavioral Health in Primary Care: A Guide for Clinical Integration. Cummings NA, Cummings JL, Johnson JN, eds. Madison, Conn: Psychosocial Press; 1997:371-398.
20. Lundin A. [Theme: High utilizers of health care services. How do we help the patients who constantly top visitors’ statistics?]. Lakartidningen. 2001;98:4320-4321.
21. CPT 2004 Physician’s Current Procedural Terminology. ed. Chicago, Ill: American Medical Association; 2003.
22. Parkinson GR, Jr., Harrell FE, Jr., Hammond WE, Wang XQ. Characteristics of adult primary care patients as predictors of future health services charges. Med Care. 2001;39:1170-1181.
23. Zhao Y, Ash AS, Ellis RP, Slaughter JP. Disease burden profiles: an emerging tool for managing managed care. Health Care Manag Sci. 2002;5:211-219.
24. Weiner JP, Starfield BH, Steinwachs DM, Mumford LM. Development and application of a population-oriented measure of ambulatory care case-mix. Med Care. 1991;29:452-472.
25. Starfield B, Weiner J, Mumford L, Steinwachs D. Ambulatory care groups: a categorization of diagnoses for research and management. Health Serv Res. 1991;26:33-74.
26. Orueta JF, Lopez-De-Munain J, Baez K, et al. Application of the national classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). 2004. Available at: http://www.cdc.gov/nchs/about/otheract/icd9/abticd9.htm. Accessed: July 28, 2004.
27. National Center for Health Statistics, Center of Disease Control. International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM). 2004. Available at: http://www.cdc.gov/nchs/about/otheract/icd9/abticd9.htm. Accessed: July 28, 2004.
28. Lemeshow S, Hosmer DW, Jr. A review of goodness of fit statistics for use in the development of logistic regression models. Am J Epidemiol. 1982;115:92-106.
29. Rothman AA, Wagner EH. Chronic illness management: what is the role of primary care? Ann Intern Med. 2003;138:256-261.
30. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 1. JAMA. 2002;288:1775-1779.
31. Bodenheimer T, Wagner EH, Grumbach K. Improving primary care for patients with chronic illness: the chronic care model, Part 2. JAMA. 2002;288:1909-1914.
32. Pearson SD, Katzenick DJ, Simon GE, et al. Depression among high utilizers of medical care. J Gen Intern Med. 1999;14:461-468.
33. Peltenburg M, Fischer JE, Bahns O, van Dulmen S, van den Brink-Muinen A. The unexpected in primary care: a multicenter study on the emergence of unvoiced patient agenda. Ann Fam Med. 2004;2:534-540.