Finding Future High-cost Cases: Comparing Prior Cost Versus Diagnosis-based Methods

Arlene S. Ash, Yang Zhao, Randall P. Ellis, and Marilyn Schlein Kramer

Objective. To examine the value of two kinds of patient-level data (cost and diagnoses) for identifying a very small subgroup of a general population with high future costs that may be mitigated with medical management.

Data Sources. The study used the MEDSTAT MarketScan® Research Database, consisting of inpatient and ambulatory health care encounter records for individuals covered by employee-sponsored benefit plans during 1997 and 1998.

Study Design. Prior cost and a diagnostic cost group (DCG) risk model were each used with 1997 data to identify 0.5-percent-sized "top groups" of people most likely to be expensive in 1998. We compared the distributions of people, cost, and diseases commonly targeted for disease management for people in the two top groups and, as a benchmark, in the full population.

Principal Findings. The prior cost– and DCG-identified top groups overlapped by only 38 percent. Each top group consisted of people with high year-two costs and high rates of diabetes, heart failure, major lung disease, and depression. The DCG top group identified people who are both somewhat more expensive ($27,292 vs. $25,981) and more likely (49.4 percent vs. 43.8 percent) than the prior-cost top group to have at least one of the diseases commonly targeted for disease management. The overlap group average cost was $46,219.

Conclusions. Diagnosis-based risk models are at least as powerful as prior cost for identifying people who will be expensive. Combined cost and diagnostic data are even more powerful and more operationally useful, especially because the diagnostic information identifies the medical problems that may be managed to achieve better outcomes and lower costs.

Key Words. Diagnostic cost group (DCG), disease management, prediction, prior cost, sensitivity, specificity

Medical costs are known to be highly concentrated, with a few people generating a large percentage of total cost in any year (Anderson and Knickman 1984; Zook and Moore 1980). For example, in the 1996 Medical Expenditure Panel Survey, the 1 percent of the population that cost the most consumed 27 percent of the

resources; the top 5 percent consumed 55 percent; and the top 10 percent consumed 69 percent (Berk and Monheit 2001). However, individual health costs have a large random component; these striking figures do not mean that it is the same few people who consistently account for the bulk of health care spending. Moreover, individuals move in and out of the high-cost group.

The ability to prospectively identify future high-cost people is important for reinsurance and other aspects of financial management. However, medical managers also need to know when expected high costs have the potential to be reduced—and coordination and quality of care enhanced—with case management.

In this article, we first examine persistence and change in expenditure levels in a large, privately insured population over two years. Then, we compare two methods for identifying 0.5 percent–sized "top groups" of people with the highest expected future costs and examine the prevalence in these groups of conditions that are well-suited for management to effect future health and utilization. While some previous research has looked at 5 percent, 10 percent, and 20 percent high-cost subgroups (Meenan, O'Keeffe-Rosetti, Hornbrook, et al. 1999), we chose the 0.5 percent group size to illustrate how a manager might identify a group both small enough and expensive enough to justify intensive case management. The 0.5 percent cut corresponds roughly to greater than $40,000 in this year's cost or more than $25,000 in next year's expected cost. A "successful" top group will: (1) have high average cost next year, (2) contain few people whose next-year costs are low, and (3) contain many people with potentially manageable diseases.

Specifically, we compared two methods for identifying top groups using year-one data. The first identifies the 0.5 percent of the population with the highest year-one total cost. The second uses a diagnostic cost group (DCG) prediction model to identify an equal number of people with the highest expected cost next year. Prior cost is traditionally used by actuaries and underwriters to identify people whose costs will be high (Bluhm and Koppel 1988; Cookson 1996). The DCG model and classification system are widely used for predicting average payments and comparing the average health status of groups (Ash, Ellis, Pope, et al. 2000),

---

Address correspondence to Arlene S. Ash, Ph.D., Research Professor, Department of General Internal Medicine, Boston University, 720 Harrison Avenue, #1108, Boston, MA 02118. Arlene S. Ash is also a Senior Scientist with D_xCG, Inc., Boston. Yang Zhao, Ph.D., is Senior Research Associate, D_xCG, Inc. Randall P. Ellis, Ph.D., is Professor, Department of Economics, Boston University, and Senior Scientist, D_xCG, Inc. Marilyn Schlein Kramer, M.B.A., is President, D_xCG, Inc. This article, submitted to *Health Services Research* on August 1, 2001, was revised and accepted for publication on October 18, 2001.

and various researchers have examined the predictive power of prior-cost and other models (Ash, Porell, Gruenberg, et al. 1989; Cohen and MacWilliam 1995; Epstein and Cumella 1988; Lamers 1999; Meenan, O'Keeffe-Rosetti, Hornbrook, et al. 1999). In early work, prior-cost models yielded higher R^2 values than all models that avoided using such data, although it now appears that more refined diagnostic models may be producing higher R^2 values than models that rely on age, sex, and costs only (Zhao et al. forthcoming). However, prior cost and diagnosis-based models have not been previously described with respect to their ability to prospectively identify small subsets of high-cost cases.

Our hypotheses for this study were:

- A diagnosis-based risk model can identify people who will be high cost next year as well as or better than prior cost.
- Diagnosis-based and prior-cost methods identify substantially different groups of potentially high-cost people.
- People predicted to be high cost by both methods will be particularly expensive.
- The diagnosis-based method in particular identifies cases with a high prevalence of diseases that are frequently targeted for case management.

METHODS

Data

We obtained data from the MEDSTAT MarketScan® Research Database, the largest multisource private sector health care database in the United States, capturing inpatient and outpatient health care service use by individuals covered by large employer-sponsored benefit plans during 1997 and 1998. More than 100 payers, covering fee-for-service, fully capitated, and partially capitated plans from all regions of the country, were included in each year. MarketScan is widely recognized by public and private researchers for its comprehensiveness and quality and has been cited frequently in peer-reviewed journal articles (e.g., Crown, Hylan, and Meneades 1998; Goodman, Nishiura, and Hankin 1998; Hillman, Joseph, Mabry, et al. 1990; Hu and Rush 1995; Leslie and Rosenheck 1999; Iezzoni 1997). We selected the approximately 2.7 million individuals eligible for at least one month in each of two study years, of whom 73 percent had fee-for-service coverage; the rest were enrolled in a range of capitated plans. Our population was 52 percent

female. Children less than age 18 accounted for 23 percent; adults aged 18 to 44 were 40 percent; and only 0.2 percent were over age 64. The mean age was 35.

The key outcome variable for this study was total medical cost in 1998, calculated for each individual by adding "covered expenses" for all inpatient and ambulatory care.[1] Covered expenses included deductibles, coinsurance, and coordination-of-benefits payments. Because only some of our population had pharmacy benefits, we did not include outpatient pharmacy costs in the total.

Identifying Top Groups

We used the DCG/hierarchical condition category (HCC) prospective model described by Ash, Ellis, Pope, et al. (2000), as implemented in $D_xCG^{®}$ Release 5.0 software, to characterize the health status of individuals. This model predicts next year's cost for each individual based on age, sex, and the range of distinct medical problems encountered this year. It uses diagnoses as recorded in International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM) codes from both inpatient and outpatient claims. Each ICD-9-CM code is classified into one of 118 condition categories (CC). Each CC encompasses similar clinical problems with similar expected costs. People can have multiple CCs; those with no medical encounters have none. Clinical hierarchies are imposed among CCs to produce HCCs, which identify for each person the most costly manifestation of each distinct disease. A person classified as belonging to one "hierarchicalized" CC cannot also belong to an HCC of lower rank in the same hierarchy. Because the HCCs represent an exhaustive classification of all medical problems recorded in diagnoses, it is not uncommon for a sick person, especially one with multiple comorbid conditions, to have ten or more HCCs noted. The DCG/HCC model used here predicts costs from the set of comorbid diseases recognized. It is not specifically designed for the purpose of identifying top groups.

We applied default cost weights included in the D_xCG software to the 1997-identified HCCs and age and sex to get individual predictions for 1998. We selected the "DCG top group" containing 0.5 percent of the population with the highest DCG predictions (n = 13,328). To make the prior-cost top group directly comparable, we included in it those 13,328 people whose 1997 costs were highest.

ANALYSES

We described the distribution of health care costs in 1998 (year two) and examined how individuals' costs changed in groups based on year-one (prior) costs: bottom 80 percent, next 10 percent, next 15 percent, next 4 percent, next 0.5 percent, and top 0.5 percent. We also compared the 1998 cost distribution for the two top groups and, as a benchmark, the whole population in four categories (< $5,000, $5,000 to $9,999, $10,000 to $24,999, and $25,000+) to see how expensive these prospectively identified people actually were.

We further explored the overlap between our two top groups, forming a pooled population with three subgroups: persons in a top group (1) by only the DCG method, (2) by only the prior-cost method, or (3) by both methods (the "overlap"). We calculated the prevalence, percent of total population expenses incurred by, and average relative costs (as compared to the population average) for each group.

Finally, we analyzed the ability of each method to identify future high-cost people who have the common diseases most widely targeted for disease management: diabetes, congestive heart failure (CHF), asthma/chronic obstructive pulmonary disease (COPD), and depression. For each of these four chronic diseases identified by the DCG's CCs, we calculated the prevalence and relative cost by dividing the mean cost of those with the medical problem by the population average. We performed the calculation for three groups: the two top groups and the full population. We also calculated the mean and median numbers of distinct medical conditions present (that is, the number of HCCs) for people in these three groups.

RESULTS

Costs in 1998 were highly skewed, with a median of $240 and mean of $1,651. Fully one-quarter had zero costs, whereas the top 0.5 percent used 23 percent of the year-two dollars; the top 1 percent used 31 percent (Table 1). Figure 1 shows both how concentrated costs are in a given year and how much less extreme are next year's costs for this year's highest- and lowest-cost people. Also, in looking backward (data not in the tables) for the 1 percent of the population that cost the most in 1998, they had used only 11 percent of the resources and only one in five of them had been in the top 1 percent in the prior year. At the same time, the

least expensive 80 percent in 1998, who had used just 12 percent of resources that year, had consumed nearly 50 percent of the previous year's resources.

Table 1: Health Care Costs in 1998*

N	2,665,678
Mean	$1,651
Standard deviation	$7,991
Coefficient of variation x 100	484
Median	$240
80th percentile	$1,354
99th percentile	$23,697
99.5th percentile	$39,064
% of population with zero costs	25.6
% of dollars used by most expensive 20%	87.1
% of dollars used by most expensive 1%	31.1
% of dollars used by most expensive 0.5%	22.9

*For persons with at least one month of eligibility in each of 1997 and 1998.

Figure 1: Distributions of Year-one and Year-two Cost by Year-one Cost Group

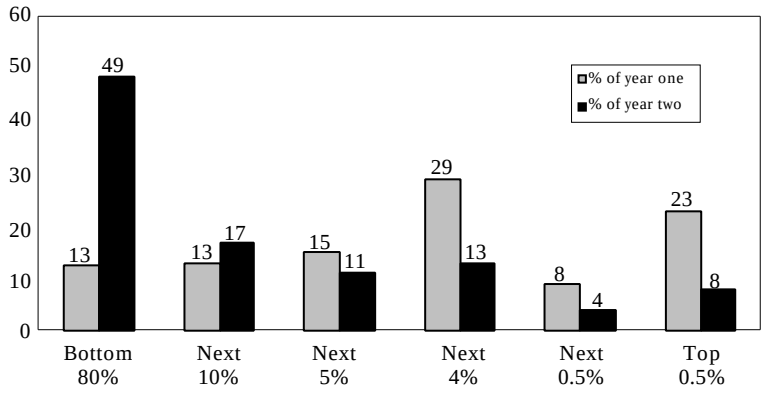


Figure 2 shows the actual 1998 cost distribution for the two top groups and the full population (as a benchmark). Less than 7 percent of the full population cost more than $5,000, 3 percent cost more than $10,000, and 0.8 percent cost more than $25,000. Both methods can identify a very small subgroup with future high-cost people; less than 47 percent of the people in either top group cost less than $5,000, and more than 41 percent cost more than $10,000 in year two. Average costs in the DCG top group were a little higher than those in the prior-cost top group ($27,292 vs. $25,981).

Figure 3 describes our top groups. There were 21,575 people (about 0.8 percent of the full population of 2.7 million) contained in at least one of the two top groups. A small group of 5,081 people (38 percent of either top group and 0.19 percent of the full population) was in the overlap. People identified by the DCG model alone were a little more expensive than those identified by prior cost alone ($16,493 vs. $14,510); people in the overlap were extraordinarily expensive ($46,219), with costs 28 times higher than average.

Table 2 shows the prevalence and relative year-two costs of four important manageable diseases by top group status. In the full population, about 3 percent had diagnoses indicating diabetes, and 0.5 percent had diagnoses indicating congestive heart failure; just 8 percent had a diagnosis for at least one of the four diseases. Patients with these manageable conditions concentrated in the top groups. Fully 28 percent and 20 percent of those in the DCG and prior-cost top groups, respectively, had diabetes; 20 percent and 18 percent had CHF; and 49 percent and 44 percent had at least one of the four diseases. Thus, people identified by the DCGs were slightly more suitable for case management than those identified by

Figure 2: Actual Year-two Cost Distribution
by Year-one Top Group Status

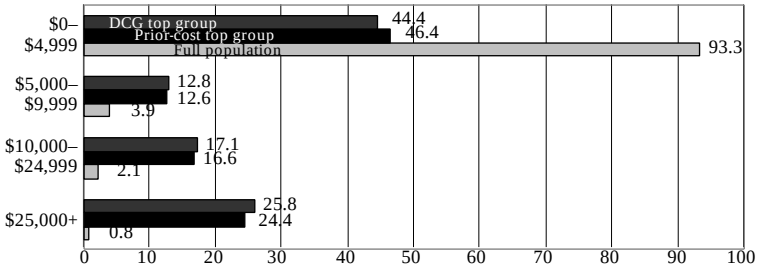
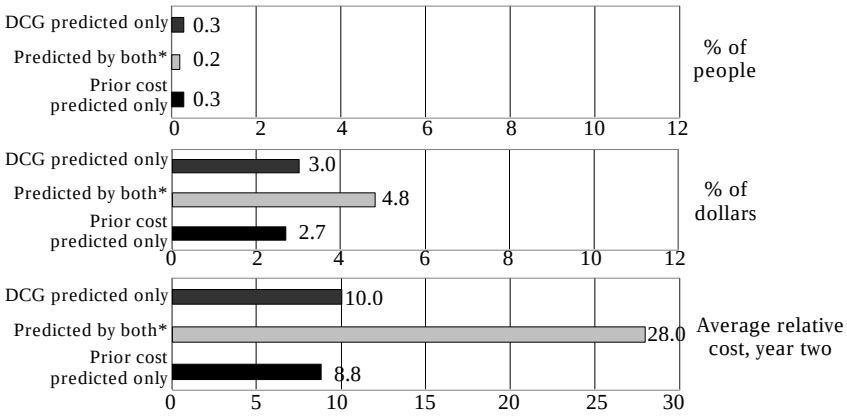


Figure 3: Top Groups as Predictors of Future Cost: DCG Vs. Prior Cost



*People identified by both the DCG and prior-cost models (the overlap).

Table 2: Year-one Prevalence and Relative Year-two Costs by Top Group Status for Disease Cohorts

	Full Population		*DCG Top Group*		*Prior-cost Top Group*	
	*Prevalence**	*Relative Cost*†	*Prevalence**	*Relative Cost*†	*Prevalence**	*Relative Cost*†
Full population	100.0	1.0	100.0	16.5	100.0	15.7
Disease cohort						
Diabetes	2.8	3.5	28.4	19.7	19.7	20.9
CHF	0.5	7.6	20.2	21.3	17.6	20.7
Asthma/COPD	3.5	2.2	15.9	16.3	15.3	17.4
Depression	2.0	2.3	4.8	15.0	6.2	15.2
Any of the above	8.2	2.7	49.4	18.6	43.8	18.4

*The percentage of the column-specified group that belongs to the row-specified cohort.
†Average cost for this subgroup divided by the average cost for the full population ($1,651).

the prior-cost model. Not explicitly shown in Table 2 is the fact that multiple co-morbidities were extremely common in the top groups and far more common than in the full population. Specifically, the median numbers of HCCs present in the full population, the prior-cost top group, and the DCG top group were, respectively, two, ten, and ten. The analogous averages were 2.5, 10.2, and 10.5.

People in each of the four disease cohorts were predictably expensive, with next year's costs ranging from 2.2 to 7.6 times the average. However, within disease-identified cohorts, those in the top groups were substantially more expensive, ranging from 15.0 to 21.3 times the average. For example, splitting the CHF cohort by DCG top group status, approximately one-fifth in that top group averaged $35,211 each and used more than 50 percent of the CHF-cohort dollars; this contrasts with the much larger group of people with CHF who were not in the DCG top group and averaged only $6,901 each. In asthma/COPD and depression, the top groups identified by either method were each between six and eight times as expensive as their disease cohort averages.

DISCUSSION

Health care expenditures for the privately insured population in the United States are highly skewed. In any given year, more than one-quarter of the population incurs either no or minimal medical costs, while the most expensive 1 percent has annual costs of more than $20,000 and absorbs approximately 30 percent of all expenditures. However, movement from inexpensive to expensive is common; we have sought to identify a manageably small subgroup of people with high costs next year. Clearly, the same methods can be used to identify a "sickest" subgroup of whatever size is desired for any cohort of interest.

Sensitivity and specificity are the most common measures of a screening mechanism's ability to correctly classify people as to the presence or absence of a specified problem. A measure is specific if most of the people who do not have the problem are classified as nonproblematic; it is sensitive if most of the people who do have the problem are classified as problematic.

However, specificity is not useful in this setting. This is because specificity cannot be less than 99 percent for any 0.5-percent-sized group used to detect an uncommon problem.[2] Moreover, sensitivity, while somewhat discriminating, is also problematic. For example, if we define "high cost" as exceeding $5,000 in year two, even if every case in our 0.5 percent group was high cost, sensitivity would only be 7 percent because the "true high cost" group is nearly 14 times larger than the 0.5 percent that the screen identifies. If true high cost were instead defined as "exceeds $25,000," a perfect screen would still miss nearly 40 percent of cases because there would be that many more problems than the 0.5 percent that are correctly identified. "Solving" this problem by setting the threshold even higher leads to $25,000 top group cases being classified as non–high cost

"errors." Alternatively, we propose examining equal-sized, model-identified top groups with respect to their entire future cost distributions and other characteristics (such as the prevalence of key diseases) to compare their ability to identify managerially relevant subgroups of high-risk people.

To "save costs and improve quality," case managers must prospectively identify likely high-cost cases that are amenable to intervention. We have previously shown that the most costly people have many comorbidities (Zhao et al. forthcoming). We have seen here that both prior-cost and DCG top groups contain many people with potentially manageable problems such as diabetes, respiratory problems, and CHF. In contrast to prior-cost data alone, the DCG model can be further used to identify individuals whose full morbidity profiles seem most promising for particular management strategies.

Starting in the early 1980s, researchers sought to develop models to predict cost that avoided using prior cost. The main reason was the desire to infer level of need from the medical problems present rather than from a variable that is heavily influenced by "practice style" in addition to medical need. In addition, diagnoses can distinguish among medical problems with the same current costs but very different future cost implications. However, it was originally feared that models that did not incorporate costs would never match the predictive power of cost. In these data, with respect to the measures considered here, the diagnosis-based method performed as well as or better than prior cost.

To identify a DCG-based top group, we applied "benchmark" formulas to age, sex, and lists of medical diagnoses for the base year (1997); no cost data for either base or target year were required. The DCG cost weights were generated from a nationally representative population. There may be some model overfitting because our analytic file is from the MarketScan database, the same data source for the benchmark data. However, the file represents the experience of 2.7 million people; in previous experiences when DCG model coefficients were determined using 1.3 million lives, fitting and validating performance measures, such as R^2, were essentially identical (Ash, Ellis, Yu, et al. 1998). In predicting outcomes for a specific disease cohort, DCG predictions recalibrated to that cohort might perform even better. Additionally, the most useful application of a diagnosis-based method for case identification would involve frequent updates of the data, so that people could be identified soon after they become at risk. We have previously shown that model performance improves dramatically when inputs to these models are updated monthly (Ellis and Ash 1989).

Identifying a future high-cost top group using prior cost requires no diagnoses, only expense records. However, individual-level medical information is still

needed if cases are to be identified and appropriately managed, and a prediction model such as the DCG is desirable for comparing the actual cost of managed patients with the expected costs in the absence of management. Clearly, managers with both diagnostic and cost data will be in the best position to plan for both medical and financial contingencies.

CONCLUSIONS

With respect to our hypotheses, we have:

- Demonstrated, in a privately insured population, that being a very high cost case in a particular year is a transient condition that only partially overlaps with "having diseases that predict high future costs" or with having been high cost last year or with being high cost next year;
- Clarified why sensitivity and specificity are poor measures of the ability of models to identify a manageably small set of future high-cost cases;
- Proposed a new way (describing the characteristics of model-identified top groups) to examine and compare the ability of models to select small groups of high-risk cases that are good candidates for case management;
- Shown that both prior cost and DCG top groups are rich in people with the kinds of chronic diseases addressed by disease managers; and
- Confirmed that prior cost, which was historically superior to diagnostic information for the purpose of predicting future costs, is no longer better than the current generation of diagnosis-based risk models for predicting future costs. In fact, the DCG/HCC model used in this study proved to be slightly superior to prior cost at identifying a top group with high costs and high prevalence of the diseases that are commonly targeted for case management.

ACKNOWLEDGMENT

We are grateful to John Z. Ayanian for his careful review of an earlier version of this article.

NOTES

1. Costs for people who were present for only part of 1998 were "annualized" and counted as fractional observations in calculating averages. For example, a person who died in June, having cost $6,000, contributes half of an observation, with annualized spending of $12,000.
2. Because 99.5 percent of people are not in the top group, if p percent have the problem, the worst we could possibly do is to have all 0.5 percent that are in the top group not be problems. In that case, specificity would equal $S = (100 - p - 0.5) / (100 - p)$; S increases with increasing p and equals 0.99 when $p = .50$.

REFERENCES

Anderson, G. F., and J. Knickman. 1984. "Patterns of Expenditures Among High Utilizers of Medical Care Services." *Medical Care* 22 (2): 143–49.

Ash, A., R. P. Ellis, G. C. Pope, J. Z. Ayanian, D. W. Bates, H. Burstin, L. I. Iezzoni, E. McKay, and W. Yu. 2000. "Using Diagnoses to Describe Populations and Predict Costs." *Health Care Financing Review* 21 (3): 7–28.

Ash, A., R. P. Ellis, W. Yu, E. A. McKay, L. I. Iezzoni, J. Z. Ayanian, D. W. Bates, H. Burstin, S. Byrne-Logan, and G. C. Pope. 1998. "Risk Adjustment for the Non-elderly." Report of the Health Care Financing Administration, Cooperative Agreement No. 18-C-90462 / 1-02.

Ash, A., F. Porell, L. Gruenberg, E. Sawitz, and A. Beiser. 1989. "Adjusting Medicare Capitation Payments Using Prior Hospitalization Data." *Health Care Financing Review* 10 (4): 17–29.

Berk, M. L., and A. C. Monheit. 2001. "The Concentration of Health Care Expenditures, Revisited." *Health Affairs* 20 (2): 9–18.

Bluhm, W. F., and S. Koppel. 1988. "Individual Health Insurance Premiums." In *Individual Health Insurance*, edited by F. T. O'Grady, 59–61. Schaumburg, IL: Society of Actuaries.

Cohen, M. M., and L. MacWilliam. 1995. "Measuring the Health of the Population." *Medical Care* 33 (Suppl. 12): DS21–42.

Cookson, J. P. 1996. "Medical Claim Cost Trend Analysis." In *Group Insurance*, edited by W. F. Bluhm, 519–47. Winsted, CT: ACTEX Publications.

Crown, W., T. Hylan, and L. Meneades. 1998. "Antidepressant Selection and Use and Healthcare Expenditures: An Empirical Approach." *Pharmacoeconomics* 13 (4): 435–48.

Ellis, R., and A. Ash. 1989. "The Continuous Update Diagnostic Cost Group Model." Health Policy Research Consortium, Report to the Health Care Financing Administration, Cooperative Agreement No. 15-C-98835 / 1-03.

Epstein, A. M., and E. J. Cumella. 1988. "Capitation Payment: Using Predictors of Medical Utilization to Adjust Rates." *Health Care Financing Review* 10 (1): 51–69.

Goodman, A. C., E. Nishiura, and J. R. Hankin. 1998. "A Discrete Choice Model of Drug Abuse Treatment Location." *Health Services Research* 33 (1): 125–45.

Hillman, B. J., C. A. Joseph, M. R. Mabry, J. H. Sunshine, S. D. Kennedy, and M. Noether. 1990. "Frequency and Costs of Diagnostic Imaging in Office Practice—A Comparison of Self-referring and Radiologist Referring Physicians." *New England Journal of Medicine* 323 (23): 1604–08.

Hu, T. W., and A. J. Rush. 1995. "Treatment Patterns and Costs of Treatment in the Private Sector of the United States." *Social Psychiatry and Psychiatric Epidemiology* 30 (5): 224–30.

Iezzoni, L. I. 1997. "Assessing Quality Using Administrative Data." *Annals of Internal Medicine* 127 (8, Part 2): 666–74.

Lamers, L. M. 1999. "Risk-adjusted Capitation Based on the Diagnostic Cost Group Model: An Empirical Evaluation with Health Survey Information." *Health Services Research* 33 (6): 1727–44.

Leslie, D. L., and R. Rosenheck. 1999. "Changes in Inpatient Mental Health Utilization and Costs in a Privately Insured Population, 1993 to 1995." *Medical Care* 37 (5): 457–68.

Meenan, R. T., M. C. O'Keeffe-Rosetti, M. C. Hornbrook, D. J. Bachman, M. J. Goodman, P. A. Fishman, and A. V. Hurtado. 1999. "The Sensitivity and Specificity of Forecasting High-cost Users of Medical Care." *Medical Care* 37 (8): 815–23.

Zhao, Y., A. Ash, R. P. Ellis, and J. P. Slaughter. Forthcoming. "Disease Burden Profiles: An Emerging Tool for Managing Managed Care." *Health Care Management Science* (in press).

Zook, C. J., and F. Moore. 1980. "High-cost Users of Medical Care." *New England Journal of Medicine* 302 (18): 996–1002.