Finding Future High Cost Cases: Comparing Prior-Cost versus Diagnoses-based Methods

Arlene Ash, Ph.D.[1, 2]
Yang Zhao, Ph.D.[2]
Randall P. Ellis, Ph.D.[3, 2]
Marilyn Schlein Kramer[2]

October 5, 2001

[1] Department of General Internal Medicine, Boston University,
[2] D_xCG, Inc., Boston, MA
[3] Department of Economics, Boston University

For correspondence contact the principal author:
Arlene Ash, Ph.D., 720 Harrison Ave., #1108, Boston University, Boston, MA 02118, Phone: 617-638-7518, Fax: 617-638-8026. Email: aash@bu.edu.

Comparing Costs and Diagnoses For Finding Future High Cost Cases

Objectives. 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. Medstat's MarketScan Research Database, consisting of inpatient and ambulatory healthcare 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 are each used with 1997 data to identify ½-of-1%-sized "Top Groups" of people most likely to be expensive in 1998. We compare 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 overlap by only 38%. Each Top Group consists of people with high year-2 costs and high rates of diabetes, heart failure, major lung disease and depression. The DCG Top Group identifies people who are both somewhat more expensive ($27,292 vs. $25,981) and more likely (49.4% vs. 43.8%) than the Prior-cost Top Group to have at least one of diseases that are commonly targeted for disease management. The overlap group average cost is $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), prediction, prior cost, disease management, 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 & Knickman 1984; Zook & Moore 1980). For example, in the 1996 Medical Expenditure Panel Survey (MEPS), the 1% of the population that cost the most consumed 27% of the resources; the top 5% consumed 55%; the top 10%, 69% (Berk & 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 healthcare spending. Moreover, individuals move in the 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 paper, 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 ½-of-1%-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 impact future health and utilization. While some previous research has looked at 5%, 10% and 20% high-cost subgroups (Meenan, et al. 1999), we chose the 0.5% 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% 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 compare two methods for identifying Top Groups using year-1 data. The first identifies the ½ of 1% of the population with the highest year-1 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 is widely used for predicting average payments and comparing the average health status of groups (Ash, et al., 2000), and various researchers have examined the predictive power of prior-cost and other models (Ash, et al., 1989; Cohen and MacWilliam 1995; Epstein and Cumella 1988; Lamers 1999; Meenan, 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 now be producing higher R^2s than models that rely on age, sex and costs only (Zhao, et al. 2001). 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 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 Medstat's MarketScan® Research Database, the largest multi-source private sector healthcare database in the U.S., capturing inpatient and outpatient healthcare service use by individuals covered by large-employer-sponsored benefit plans during 1997 and 1998. More than 100 payers, covering fee-for-service, fully- 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 et al 1990; Hu and Rush 1995; Leslie and Rosenheck 1999; Iezzoni 1997). We selected the approximately 2.7 million individuals eligible at least one month in each of two study years, of whom 73% had fee-for-service coverage and the rest were enrolled in a range of capitated plans. Our population is 52% female. Children less than age 18 account for 23%; adults aged 18 to 44 are 40%; only 0.2% are over age 64. The mean age is 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.[2] 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/HCC prospective model described in Ash, 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 (CCs). 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 Hierarchical Condition Categories (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 ½ of 1% 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 2) and examined how individuals' costs changed in groups based on year-1 (prior) costs: Bottom 80%, Next 10%, Next 15%, Next 4%, Next ½ of 1%, and Top ½ of 1%.

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) only by the DCG method, 2) only by the Prior-cost method, or 3) by both methods (the "overlap"). We calculated the prevalence, the percent of total population expenses incurred by, and the 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 the relative cost, as calculated 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 are highly skewed, with a median of $240 and mean of $1,651. Fully one-quarter has zero costs while the top ½ of 1% uses 23% of the year-2 dollars; the top 1% uses 31% (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 backwards (data not in the tables) for the 1% of the population that cost the most in 1998, they had used only 11% of the resources and only 1 in 5 of them had been in the top 1% in the prior year. At the same time, the least expensive 80% in 1998, who had used just 12% of resources that year, had consumed nearly 50% of the previous year's resources.

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

Figure 3 describes our Top Groups. There are 21,575 people (about 0.8% 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% of either Top Group and 0.19% of the full population) is in the overlap. People identified by the DCG model alone are a little more expensive than those identified by prior cost alone ($16,493 vs. $14,510); people in the overlap are extraordinarily expensive ($46,219), with costs 28 times higher than average.

Table 2 shows the prevalence and relative year-2 costs of four important manageable diseases by Top Group status. In the full population, about 3% have diagnoses indicating diabetes, and ½ of 1% indicating congestive heart failure; just 8% have a diagnosis for at least one of the four diseases. Patients with these manageable conditions concentrated in the top groups. Fully 28% and 20% of those in the DCG- and Prior-cost Top Groups, respectively, have diabetes; 20% and 21%, have CHF; and, 49% and 44% have at least one of the 4 diseases. Thus, people identified by the DCGs are slightly more suitable for case management than those identify by prior-cost model. Not explicitly shown in Table 2 is the fact that multiple comorbidities are 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 are, respectively, 2, 10 and 10. The analogous averages are 2.5, 10.2 and 10.5.

People in each of the 4 disease cohorts are predictably expensive, with next year's costs ranging from 2.2 to 7.6 times average. However, within disease-identified cohorts, those in the Top Groups are substantially more expensive, ranging from 15.0 to 21.3 times average. For example, splitting the CHF cohort by DCG Top Group status, approximately one-fifth in that Top Group average $35,211 each and use more than 50% of the CHF-cohort dollars; this contrasts with the much larger group of people with CHF who are not in the DCG Top Group and average only $6,901 each. In asthma/COPD and depression, the Top Groups identified by either method are each between 6 and 8 times as expensive as their disease cohort averages.

DISCUSSION

Health care expenditures for the privately insured population in the U.S. are highly skewed. In any given year, over one-quarter of the population incurs either no or minimal medical costs, while the most expensive 1% has annual costs over $20,000, and absorbs approximately 30% 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 non-problematic; it is sensitive if most of the people who do have the problem are classified as problematic.

However, specificity is not useful here, since it is mathematically impossible for it to be less than 99% for any ½%-sized group used to detect any uncommon problem.[3] Sensitivity, while somewhat discriminating, is still problematic in this setting. For example, if we define "high cost" as exceeding $5,000 in year 2, even if every case in our ½% group was high cost, sensitivity would only be 7%, because the "true high cost" group is nearly 14 times larger than the ½% that the screen identifies. If "true high cost" were instead defined as "exceeds $25,000," a perfect screen would still miss nearly 40% of cases, because there would be that many more problems than the ½% that are correctly identified. If we try to solve this problem by setting the threshold for a case higher than $25,000, this causes a $25,000 case to be labeled as a "failure."

In our opinion, the problem with both sensitivity and specificity is that what they measure is uninteresting in the setting of our problem. We propose comparing the costs and other characteristics of model-identified Top Groups as a natural and appropriate way to measure the ability of models to identify small subgroups of high-risk people.

Expensive patients even if slightly below a high cost threshold (e.g., $23,000 in a year and not $25,000) are clinically and operationally meaningful for disease management. "Saving costs while improving quality" is the mantra of disease management programs. However, these programs are most likely to be successful when they can routinely and inexpensively target people with the medical problems that match the intended interventions. We have seen elsewhere

that high expenses are usually due to multiple comorbidities (Zhao, et al. 2001). We have shown here that both prior-cost and DCG Top Groups contain many people with diabetes, respiratory problems and CHF, and people with multiple comorbidities. In fact, we were surprised to find that disease prevalence was only a little less in the prior-cost Top Group than in the DCG Top Group. However, the ability to identify a disease resides solely in the diagnoses; it cannot be inferred from cost data alone.

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 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 (here 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, since our analytic file is from MarketScen® 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 the R^2, were essentially identical (Ash, 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.

CONCLUSION

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.
- 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 paper proved to be slightly superior to prior cost at

identifying a Top Group with high costs and with high prevalence of the diseases that are commonly targeted for case management.

ENDNOTES

1. All the numbers presented hear are based on our analysis using MarketScan Research Database 1997-1998.
2. 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 ½ of an observation with annualized spending of $12,000.
3. Since 99.5% of people are not in the Top Group, if tg percent have the problem, the worst we could possibly do is to have all 0.5% that are in the Top Group not be problems. In that case, specificity would equal $S = (100 - tg - 0.5)/(100 – tg)$; S increases with increasing tg and equals 0.99 when $tg = 0.50$.

REFERENCES

Anderson, G. F., and J. Knickman. 1984. “Patterns of Expenditures among High Utilizers of Medical Care Services.” Medical Care, 22 (2): 143-149.

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.

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. “Risk Adjustment for the Non-Elderly.” Health Care Financing Administration, Cooperative Agreement No. 18-C-90462/1-02 (1998).

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.

Bluhm W. F. and S. Koppel. 1988. “Individual Health Insurance Premiums.” In Individual Health Insurance, edited by F. T. O’Grady, pp.59-61. Schaumberg, Illinois: Society of Actuaries.

Berk, M.L. and A.C. Monheit. 2001. “The Concentration of Health Care Expenditures, Revisited.” Health Affairs, March/April: 9-18.

Cohen, M. M., and L. MacWilliam. 1995. “Measuring the Health of the Population.” Medical Care, 33 (12 Suppl): DS21-42.

Cookson J. P. 1996. “Medical Claim Cost Trend Analysis.” In Group Insurance, edited by W. F. Bluhm, pp.519-547. 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): 435-448.

R Ellis and A. Ash. "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. June 1989.

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: 125-145.

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): 1604-1608.

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): 224-230.

Iezzoni, L. I. 1997. “Assessing Quality Using Administrative Data.” Annals of Internal Medicine 127 (8 part 2): 666-674.

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): 457-468.

Meenan, R. T., M. C. O’Keefee-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. (In press.) "Disease Burden Profiles: An Emerging Tool for Managing Managed Care." *Health Care Management Science*.

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

Table 1: Health Care Costs in 1998*

N	2,665,678
Mean	$1,651
Standard deviation	$7,991
Coefficient of Variation x100	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 1/2%	22.9

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

Table 2: Year-1 Prevalence and Relative Year-2 Costs by Top Group Status and Disease Cohort

	Full Population		DCG Top Group		Prior-Cost Top Group	
	Prevalence*	**elative Cost****	**Prevalence***	**elative Cost****	**Prevalence***	**elative 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 who belong to the row-specified cohort
** Average cost for this subgroup divided by the average cost for the full population ($1,651)

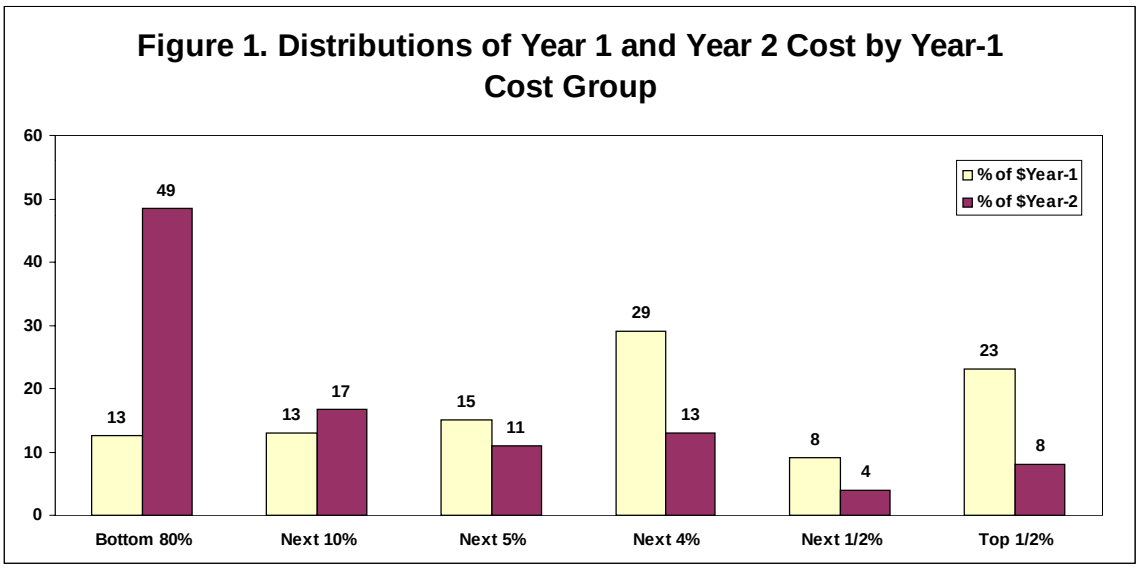


Figure 1. Distributions of Year 1 and Year 2 Cost by Year-1 Cost Group
60
50
40
30
20
10
0
% of $Year-1
% of $Year-2
13
49
13
17
15
11
29
13
8
4
23
8
Bottom 80%
Next 10%
Next 5%
Next 4%
Next 1/2%
Top 1/2%

Figure 2. Actual Year-2 Cost Distribution by Year-1 Top Group Status

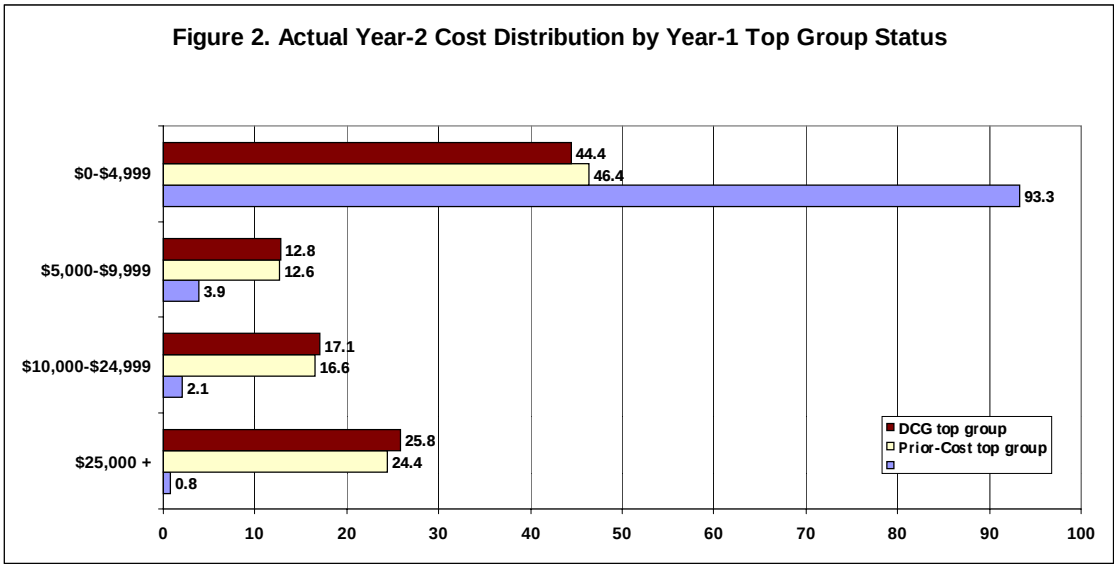
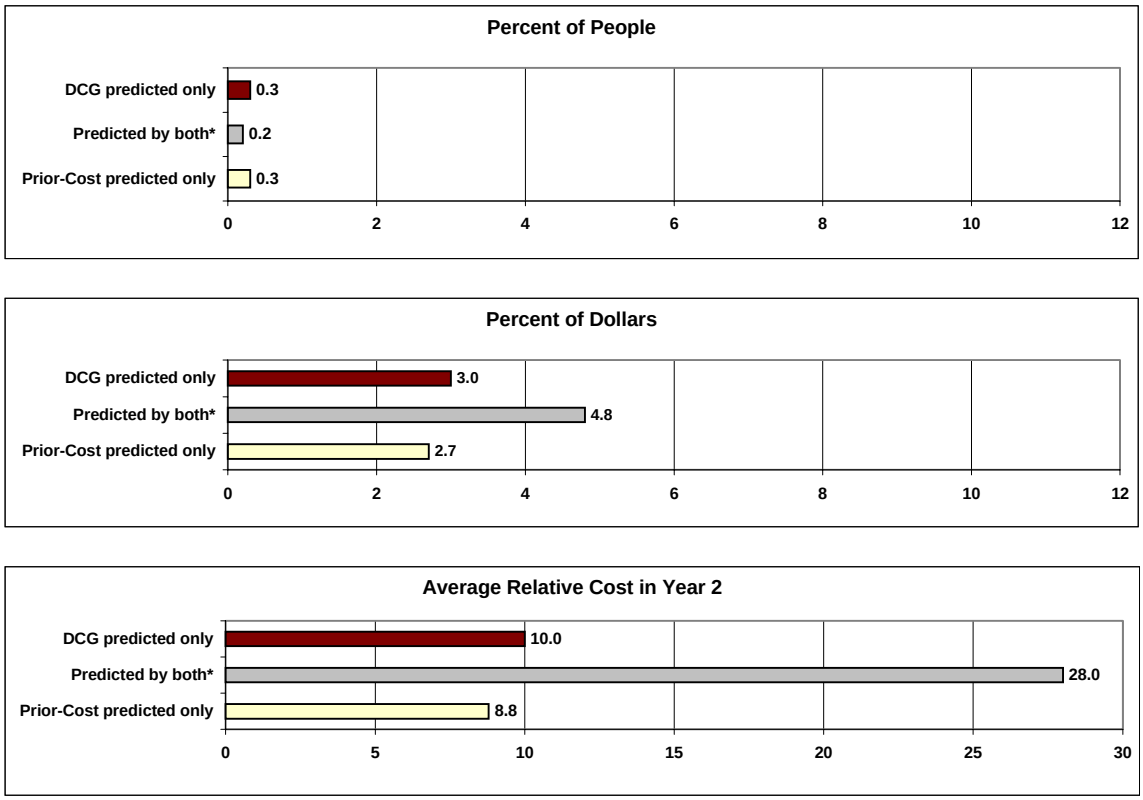


Figure 3: Top Groups as Predictors of Future Cost: DCG vs. Prior Cost



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