

The Validity of Claims-Based Risk Estimation in Underinsured Populations

Janice M. Moore, MSW; and Jean P. Hall, PhD

Objectives: To demonstrate a threat to validity in using claims-based risk tools with chronically ill, underinsured populations.

Study Design: We tracked disease burden of high-risk pool beneficiaries with potentially disabling health conditions receiving enhanced health insurance benefits through a federally funded research demonstration. At baseline, beneficiaries paid high premiums and cost sharing for risk pool coverage, and most met common criteria for underinsurance. Study benefits provided intervention group members premium and cost-sharing subsidies and additional coverage; control group members paid usual premiums and coinsurance and received usual benefits. We hypothesized that enhanced benefits for the intervention group would increase or stabilize health status measures and decrease case-mix weights, reflecting stabilized or reduced disease burden.

Methods: The SF-12v2 health survey was used to measure health status and the Johns Hopkins Adjusted Clinical Groups (ACGs), Version 8.2 with DX-PM model and prior cost for a non-elderly population, was used to measure disease burden.

Findings: Over a 3-year period, SF-12v2 scores showed stable health status for the intervention group and significant decline for the control group, while ACG case-mix weights, major illnesses, and chronic condition counts rose significantly for the intervention group but remained stable for the control group. Increased resource utilization for the intervention group appears to have driven increases in ACG measures.

Conclusions: When high cost-sharing constrains access to care, risk tools that rely on medical claims may not provide an accurate measure of disease burden.

(Am J Manag Care. 2012;18(12):e468-e476)

For author information and disclosures, see end of text.

A number of administrative tools have been developed to measure comorbidity—or burden of illness—through claims data. Instruments such as the Diagnostic Cost Group/Hierarchical Condition Category (DCG/HCC),¹ the Johns Hopkins Adjusted Clinical Groups (ACGs),² and the Chronic Illness and Disability Payment System (CDPS),³ were created to measure actuarial risk but have been extended to numerous other purposes, such as targeting high-risk patients for case management, profiling physician performance, and conducting outcomes research.⁴ Most recently, the Affordable Care Act (ACA) proposes to use individual risk scores to determine or validate average plan risk within health insurance exchanges.^{5,6} Plan average risk scores, in turn, will be used to reallocate premiums within exchanges so that plans with a disproportionate share of sicker individuals will be compensated for losses. Thus, the issue of the validity of commonly used risk tools is assuming increased importance and examining potential vulnerabilities is an urgent issue.

In the absence of readily available standardized clinical information, administrative data are often the only practical source of information for estimating illness burden. Typically tools such as the ACG and CDPS calculate a case-mix score by combining diagnostic and/or prescription drug codes with demographic variables associated with variation in utilization, such as gender and age. Some tools use prescription drug codes instead of or in addition to diagnostic codes.

Various authors have noted the limitations of these tools.⁷⁻⁹ This article makes the case for another common, unrecognized vulnerability. Specifically, the validity of administrative data as a proxy for comorbidity requires that beneficiaries have adequate access to healthcare services in order to generate a reasonably complete diagnostic profile. As this paper will demonstrate, scant or missing diagnosis codes and cost data can severely bias case-mix scores, even for very sick populations. Arguably, tools that use pharmacy data may compensate somewhat, because many chronically ill individuals use prescriptions even when they do not regularly see a provider. However, these tools are limited by the lack of specificity in the uses of many drugs. In addition, because many individuals now obtain prescriptions from \$4 generic drug retail programs, insurer prescription records can no longer be assumed to represent reliable records of utilization.¹⁰ We illustrate these vulnerabilities with the ACG by showing

In this article
Take-Away Points / e469
Published as a Web exclusive
www.ajmc.com

how scores were influenced by a dramatic increase in utilization among a group of chronically ill adults participating in a research demonstration.

METHODS

Background

The Demonstration to Maintain Independence and Employment (DMIE) was a study of disability outcomes sponsored by the Centers for Medicare & Medicaid Services (CMS) under the Ticket to Work and Work Incentives Improvement Act (TW-WIIA) of 1999 and independently conducted by 4 states between 2006 and 2009.¹¹ The demonstrations tested the hypothesis that improved access to healthcare and personal supports would prevent or delay transition to reliance on Social Security disability benefits among a population with potentially disabling conditions. The intervention was targeted to uninsured or underinsured populations whose high healthcare costs may have resulted in their delaying or forgoing needed health services. We hypothesized that the intervention would work over time by reducing barriers to care seeking, thus improving health status and reducing the burden of illness. This paper describes the experience of the Kansas DMIE, which recruited participants from the state high-risk insurance pool. The case-mix findings reported below pertain to the subset of the study population that was continuously enrolled for the first 2 years of the study. A more complete description of the full study population and outcomes is reported elsewhere.^{12,13}

As with most other state high-risk pools, the Kansas pool offers a benefit package modeled on nongroup coverage, which typically has less generous benefits and is more expensive than group coverage. At baseline, premiums were 125% of standard market rates. Deductibles ranged from \$500 to \$7500, with no benefits available prior to meeting the deductible, other than a \$250 preventive benefit in some plans. Prescription coinsurance was 50%, and medical coinsurance was 30% in-network and 50% out-of-network, with out-of-network charges not accruing toward the annual coinsurance maximum. Some plans had a \$100,000 annual maximum, and all had a lifetime limit of \$1 million. The coinsurance cap for in-network services was \$5000 for single policies and \$14,000 for family plans, with some plans having unlimited 10% coinsurance thereafter.

Study Design

The study design was a clinical trial with applicants randomly assigned to equal-sized intervention and control

Take-Away Points

- Risk tools based on *International Classification of Diseases, Ninth Revision, Clinical Modification* codes assume that access to care is sufficient to adequately represent health status.
- Access barriers may mean that many conditions, even major comorbidities, are not coded.
- While the Affordable Care Act (ACA) may bring coverage to millions of previously uninsured individuals, many will still face high cost sharing. In addition, they may have higher than average comorbidity because of lower socioeconomic status.
- Since the ACA proposes to use individual risk scores in adjusting risk across exchanges, and because claims-based risk tools are widely used for other purposes, the potential bias associated with underutilization should be recognized.

groups. Control group members received high-risk pool benefits as usual, plus cash stipends for participation in surveys (\$1200 paid in progressively increasing installments) and focus groups (\$25 per session).

The intervention consisted of premiums subsidized to a flat \$152 per month; elimination of all deductibles and coinsurance; cost-sharing limited to a \$3 copay per service and charges in excess of reasonable and customary for out-of-network services; and a \$1 million lifetime limit for DMIE-paid services. The intervention also provided nursing case management to coordinate care and pre-authorize benefits.

To be eligible, participants had to be enrolled in the state high-risk pool at least 6 months prior to recruitment and be aged 18 to 60 years, so that no one would turn 65 years during the study and attain Medicare coverage. They had to experience 1 or more health conditions that represented a potential for disability, as designated by CMS and the state program administrators. Data collected included claims files, surveys including standardized health status assessments and other questions, and focus group discussions. This research design and evaluation plan was approved by the institutional review board affiliated with the authors.

Participant Recruitment

The program began recruiting participants in the fall of 2005 and offered intervention services from April 2006 through September 2009. Over this time a total of $n = 508$ subjects participated. The present analysis is limited to the subset of subjects who were continuously enrolled during the first 2 years of the study ($n = 317$, cohorts 1 and 2). The recruitment process is described in more detail elsewhere.^{12,13}

Although intervention and control groups were equal sized at baseline, large control group attrition resulting from dropping high-risk pool coverage created an imbalance in group size over time ($n = 184$ intervention, $n = 133$ control by the end of the second year). Comparison of demographic and health status data, including interviews that tracked reasons for disenrolling and types of replacement coverage, showed

■ METHODS ■

■ **Table 1.** DMIE Participant Demographics by Group and Total at Baseline

Characteristic	Intervention	Control	Total
N	184	133	317
Female	48.9%	50.4%	49.5%
Age, y, mean (SD)	51.3 (8.2)	51.8 (8.7)	51.5 (8.4)
18-39	7.1%	7.5%	7.3%
40-49	26.1%	25.6%	25.9%
50-59	55.4%	49.6%	53.0%
60-61 ^a	11.4%	17.3%	13.9%
Educational attainment			
More than 4-year college degree	19.0%	21.1%	19.9%
Some college or 4-year degree	59.8%	59.4%	59.6%
High school diploma or less	21.2%	19.5%	20.5%
Marital status			
Married	55.7%	58.6%	57.0%
Single	22.4%	18.0%	20.6%
Divorced	19.7%	19.5%	19.6%
Widowed or separated	2.2%	3.0%	2.5%
Homeowner	80.9%	88.7%	84.2%
Own income, \$, mean (SD)^b	47,698 (58,716)	53,947 (73,882)	50,339 (65,514)
Median, \$	30,000	33,000	30,000
Range	350,000	600,000	600,000
Household annual income, \$, (SD)^b	70,790 (72,518)	72,069 (78,328)	71,333 (74,918)
Median	50,000	46,500	50,000
Range	600,000	500,000	600,000
Household income as % of federal poverty levels^b			
<133%	12.5%	10.0%	11.4%
133%-199%	9.7%	10.0%	9.8%
200%-299%	14.2%	13.1%	13.7%
300%-399%	21.6%	23.1%	22.2%
≥400%	42.0%	43.8%	42.8%
Plan deductible			
\$500	2.2%	3.0%	2.5%
\$1000	19.6%	15.8%	18.0%
\$1500	21.7%	24.1%	22.7%
\$2500	10.3%	10.5%	10.4%
\$5000	27.2%	29.3%	28.1%
\$7500	19.0%	17.3%	18.3%
Premium, \$, mean (range)	\$456 (\$822)	\$465 (\$657)	\$460 (\$822)
Held medical debt^c	19.6%	26.3%	22.4%

DMIE indicates Demonstration to Maintain Independence and Employment; SD, standard deviation.

All between-group differences are statistically non-significant.

^aAlthough the eligibility cap was set at 60 years of age, 2 individuals turned 61 years between enrollment and the beginning of the intervention.

^bFour individuals (3.5%) refused to report income at baseline.

^cQuestion was asked approximately 8 months following baseline.

that those who disenrolled did not differ significantly from those who remained. Most disenrollees obtained group coverage through their own or a spouse's employment, where there was no underwriting for pre-existing conditions. The reason cited for leaving high-risk pool coverage was almost exclusively the unaffordability of coverage.

Sample Baseline Characteristics

Study participants were demographically similar to the overall high-risk pool population, except that they were on average slightly younger, reflecting the study's eligibility cap. We found no statistically significant between-group differences in demographic variables or distribution of comorbidities (Tables 1, 2). Most participants were aged 50 to 60 years, married, well-educated, and owned homes. Mean incomes were approximately \$50,000 for individuals and \$71,000 for families, with median individual incomes of \$30,000 and family incomes of \$50,000; 57% were below 400% of poverty when family size was included and 21% had incomes below 200% of federal poverty level. Occupations included home-based service or manufacturing enterprises, farming, professional services, and small business ownership. About 70% of participants earned at least some income from self-employment, explaining this population's lack of access to group health insurance.

Baseline claims and self-reported diagnoses showed that participants experienced numerous comorbid conditions (Table 2). Musculoskeletal pain, cardiovascular conditions, and diabetes were among the most common; 19% had cancer, either active or by history. Forty-five percent reported at least 1 activity of daily living (ADL) limitation and 46% at least 1 instrumental activity of daily living (IADL) limitation, most commonly those requiring physical strength and mobility. Based on self-reported height and weight, 75% were overweight to obese (body mass index [BMI] >25), including 32% obese (BMI >30-40) and 11% morbidly obese (BMI >40).

Baseline Health Cost Burden

Baseline claims showed high cost burden. Premiums averaged \$460 per month, but ranged from \$189 to \$1011. Premiums for those aged 50 to 59 years (53% of the sample) ranged from \$264 per month for a \$7500 deductible plan to \$878 per month for a \$500 deductible plan. Only 3% of the sample enrolled in the \$500 deductible plan, while almost one-half (46%) enrolled in \$5000 or \$7500 deductible plans. Median and mean out-of-pocket expenses were \$1639 and \$2472, respectively, compared with \$960 and \$575 nationally for non-elderly individuals covered by individual and group plans.¹⁴ Because of high deductibles and cost-sharing, participants, on average, paid 61% of allowed costs. Only 58% received

reimbursement for more than one-half of allowed costs, and one-third received no reimbursement because they did not meet their deductible. Using a common definition of underinsurance—a deductible exceeding 5% or medical expenses exceeding 10% of family income; or if low-income, medical expenses exceeding 5% of family income—82% of the sample was underinsured.¹⁵ Because our data do not include expenses for other family members, the underinsurance rate is likely even higher.

In focus group discussions and during survey interviews, participants described how medical spending caused them considerable financial hardship and emotional stress. Twenty-two percent reported medical debt, with 25% owing \$5000 or more and 6% \$20,000 or more. They told of delaying or forgoing preventive services, diagnostic testing, elective surgeries, and medical equipment and employing strategies to stretch medication supplies, such as skipping pills, selectively not filling prescriptions, and continuing to use old medications after new ones had been prescribed.¹³ They also reported enrolling in plans with high deductibles because these were the only ones for which they could afford the premiums.¹⁶

Health Status Measures. The primary instrument for measuring health status was the SF-12v2, a scale derived from the SF-36.¹⁷ We also administered the World Health Organization Quality of Life brief form (WHOQOL-BREF), the World Health Organization Health and Work Performance Questionnaires (HPQs), standardized ADL-IADL measures,¹⁸ and asked for self-reported diagnoses and health conditions. All instruments were administered by telephone at baseline and at 8-month intervals over the 41-month study period (5 rounds). SF-12v2 data are reported below for the first 4 rounds (baseline plus 3) corresponding with the first 2 years of the study. Claims data for case mix analysis are summarized for the corresponding 2-year period.

To measure comorbidity burden (case mix), we used the ACG, Version 8.2 with DX-PM model and prior cost for a non-elderly population.¹⁹ We did not include prescription data because a preliminary analysis indicated it did not significantly affect outcomes for this study population. The ACG maps each beneficiary's age, gender, and diagnostic codes to a single actuarial cell, or ACG, representing their estimated resource use. ACG reference unscaled concurrent weights, or case-mix weights, are ratios comparing estimated resource use for a given ACG with the average resource use of a nationally representative sample of insured individuals. The average case-mix weight for the national population is set at 1. Thus, someone with a case-mix weight of 2 has an estimated resource use twice the national average. Although ACG case-mix weights are measures of estimated resource use, they are also proxy measures of disease burden. Gener-

■ METHODS ■

■ **Table 2.** Major Potentially Disabling Conditions, Functional Limitations, and BMI Frequencies Within Group and Total

	Claims		Self-Report		Maximum Combined Claims and Self-Report		
	Intervention	Control	Intervention	Control	Intervention	Control	Total
ICD-9-CM Category (Codes)^a							
Immune (042, 279)	2.2	1.5	2.2	1.5	2.7	1.5	2.2
Cancers (140-165, 170-172, 174-177, 179-208, 230-235)	14.1	12.8	15.2	15.0	19.6	17.3	18.6
Diabetes/endocrine (250, 277)	28.3	23.3	27.7	22.6	30.4	27.1	29.0
Blood (282-289)	6.5	6.8	2.2	3.0	7.6	7.5	7.6
Psychiatric (294-301, 310-311)	14.1	21.1	33.2	33.8	34.2	39.1	36.3
Neurological (331-337, 340-345, 350-359)	11.4	12.8	12.5	9.8	17.4	17.3	17.4
Stroke (430-438)	2.7	2.3	1.1	1.5	2.7	3.0	2.8
Sensory (360-365, 369, 386-388)	7.1	6.8	3.8	2.3	10.3	8.3	9.5
Cardiovascular (393-398, 410-417, 420-429, 440-448)	22.3	27.8	23.4	30.1	29.3	36.8	32.5
Respiratory (491-496, 500-504)	8.7	6.8	18.5	20.3	20.1	22.6	21.1
Gastrointestinal (555-556, 570-573)	3.3	6.8	7.1	9.8	8.2	12.8	10.1
Renal (580-589)	1.6	1.5	1.6	2.3	3.3	3.0	3.2
Musculoskeletal (710-725, 731-732, 737, 741)	35.3	37.6	36.4	36.8	52.7	51.9	52.4
Limited ADL^a							
Any ADL	—	—	46.7	42.1	—	—	44.8
Walking	—	—	39.7	35.3	—	—	37.9
Transferring from bed and chair	—	—	20.7	26.3	—	—	23.0
Getting outside	—	—	14.8	13.6	—	—	14.3
Dressing	—	—	12.5	9.8	—	—	11.4
Bathing	—	—	12.5	7.5	—	—	10.4
Toileting	—	—	9.8	6.1	—	—	8.3
Eating	—	—	6.0	4.5	—	—	5.4
Limited IADL^a							
Any IADL	—	—	47.8	42.9	—	—	45.7
Heavy housework	—	—	43.7	36.4	—	—	40.6
Shopping	—	—	18.5	14.4	—	—	16.8
Travel outside walking distance	—	—	15.8	8.3	—	—	12.6
Light housework	—	—	9.2	8.3	—	—	8.8
Money management	—	—	7.1	2.3	—	—	5.1
Rx management	—	—	6.0	3.0	—	—	4.7
Meal preparation	—	—	3.8	4.5	—	—	4.1
Telephone	—	—	2.2	0	—	—	1.3
BMI^b							
Total overweight or obese	—	—	74.3	75.9	—	—	75.0
Underweight (<18.5)	—	—	2.2	1.5	—	—	1.9
Normal weight (18.5-25)	—	—	23.5	22.6	—	—	23.1
Overweight (25-30)	—	—	30.1	33.8	—	—	31.6
Obese (30-40)	—	—	32.2	31.6	—	—	32.0
Morbidly obese (40+)	—	—	12.0	10.5	—	—	11.4

ADL indicates activities of daily living; BMI, body mass index; IADL, instrumental activities of daily living; ICD-9-CM, *International Classification of Diseases, Ninth Revision, Clinical Modification*; Rx, script.

All between-group differences are statistically non-significant. N = 317.

^aDoes not total to 100% because of comorbidities wherein individuals may report more than 1 of these categories.

^bCalculated from self-reported height and weight.

ally, weights greater than 1.0 indicate a population is sicker than the national average, and weights of less than 1.0 indicate it is healthier.¹⁹

ACG software also produces other measures of disease burden, including Aggregated Diagnostic Groups (ADGs), Expanded Diagnostic Clusters (EDCs), counts of major EDCs (chronic condition counts), and resource utilization bands (RUBs). Recent versions have also included a predictive model algorithm for forecasting cost.

Our hypothesis was that SF-12v2 PCS and MCS scores would stabilize or decrease more slowly for the intervention than the control group, while ACG case-mix weights would steadily decrease for the intervention group, reflecting reduced disease burden and improved health.

FINDINGS

The data presented here were collected as part of a randomized controlled study and reflect the subsample continuously enrolled for the first 2 years of the study. For the entire study population, over 32 months, intervention group SF-12v2 PCS scores remained relatively stable but control group scores significantly declined.¹² For this subsample, the same trend was evident, even though the between-group difference did not approach statistical significance by the end of year 2. However, the within-group difference showed that the control group was declining at a significantly greater rate ($P = .01$ control group, $P = .58$ intervention, repeated measures analysis of variance [ANOVA], **Table 3**). PCS scores did not differ significantly by economic status, as measured by family income as a percent of federal poverty level.

Previous research on the Medical Outcomes Study, from which the SF-12v2 is derived, found that the average age-related decline for healthy populations 45 to 64 years of age is 0.4 points per year.²⁰ As Table 3 shows, PCS scores for the intervention group fluctuated plus or minus 1 point while control group scores declined by 2.45 points over a 24-month period. Survey data also showed different self-reported health status. At study end, 34% of the intervention group indicated their health had improved compared with 21% of the control group; conversely, only 19% of the intervention group reported worsening health compared with 31% of controls; approximately equal proportions (46% in-

■ **Table 3.** Within-Group Changes in SF-12v2 PCS Scores^a Over a 2-Year Period

	Baseline	8 Mos	16 Mos	24 Mos	P
Intervention					
Means	42.9	42.8	42.4	42.0	.50
Confidence intervals	41.1-44.7	41.1-44.6	40.7-44.1	40.2-43.9	
Control					
Means	44.4	43.0	43.1	41.9	.006
Confidence intervals	42.4-46.3	41.0-45.0	41.1-45.1	39.9-44.0	

^aThe national norm for SF-12v2 PCS scores is 50.
P = repeated measures analysis of variance.

tervention vs 48% control) said their health had remained the same ($P = .01$, Pearson χ^2).

In contrast, the trend for ACG case mix moved inversely (**Table 4**). At baseline, the groups did not differ significantly on any ACG measure. Both had case-mix weights approximately 3 times the national average (2.97 intervention vs 3.59 control). Both groups had 1.1 major illnesses (ADG) and 2.6 chronic conditions. Average costs for the control group were slightly, but not significantly, higher (\$11,124 vs \$8563).

By the end of year 1, intervention group case-mix weights had increased to 4.2 and by year 2 to 4.98 ($P = .00$, repeated measures ANOVA). Over the 2 years the percent of participants with high to very high resource use also increased from approximately one-third (31%) to one-half (52%) and average cost almost doubled, from \$8563 to \$16,725 ($P = .00$). On the other hand, there were no significant changes in any ACG metric within the control group.

As **Table 5** demonstrates, a large shift to higher acuity for intervention group members occurred in years 1 and 2. Because ACGs are mutually exclusive actuarial cells, the proportion of beneficiaries with higher-acuity ACGs displaced those with lower-acuity ACGs. The proportions of intervention members with 10 or more diagnosis clusters grew 72% the first year and almost doubled by the second year, and the proportion with 6 to 9 diagnostic clusters grew 62% during the first year.

DISCUSSION

The comparative spike in case mix rates for the group that gained greater access to services illustrates an important limitation of risk tools based on administrative data. Greater utilization of care increases the chance that more conditions will be diagnosed, even those that are only suspected. More *International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9-CM)* codes drive up case mix, regardless of whether health status has changed. For our subjects, SF-12v2 PCS scores suggest that, over the first 24 months,

■ **Table 4.** Within-Group Changes in ACG Case-Mix and Other Comorbidity Measures Over a 2-Year Period

Measure	Intervention				Control			
	Baseline	Year 1	Year 2	P	Baseline	Year 1	Year 2	P
Unscaled concurrent weight (case mix)								
Means	2.97	4.20	4.98	.00	3.59	3.10	3.07	.48
Confidence intervals	2.41-3.53	3.56-4.84	4.15-5.81		2.62-4.57	2.28-3.92	2.28-3.86	
Major ADGs								
Means	1.11	1.44	1.51	.00	1.14	1.11	1.05	.70
Confidence intervals	0.96-1.27	1.30-1.59	1.35-1.68		0.94-1.33	0.92-1.29	0.88-1.23	
Chronic condition count								
Means	2.56	3.22	3.39	.00	2.61	2.44	2.57	.56
Confidence intervals	2.27-2.85	2.91-3.54	3.03-3.75		2.25-2.97	2.05-2.82	2.23-2.92	
Average cost								
Means	\$8563	\$13,361	\$16,725	.00	\$11,124	\$10,431	\$9942	.92
Confidence intervals	6376-10,750	10,638-16,084	13,013-20,438		7276-14,973	4396-16,467	6826-13,057	
Percent with high to very high resource use								
Means	31.0	45.7	52.2	.00	30.8	25.6	30.1	.45
Confidence intervals	24.2-37.7	38.4-52.9	44.9-59.5		22.9-38.8	18.1-33.1	22.2-38.0	

ADG indicates adjusted diagnostic groups or diagnostic group.
P = repeated measures analysis of variance.

the intervention group actually experienced better health status than the control group, or, at the very least, certainly no worse. Case-mix and other ACG scores, on the other hand, suggest the intervention group experienced a dramatic worsening of health status. The same measures for controls, which differed only in the level of services received, showed no significant change. Other studies similarly have found risk scores more dependent on utilization patterns than comorbidity. For instance, Song et al found that when Medicare beneficiaries from low-intensity practice regions moved to higher practice-intensity regions, their HCC risk scores increased comparatively more than those who moved to lower practice-intensity regions.²¹ Welch et al found that case-mix scores did not predict mortality. In fact, the case fatality rate among Medicare beneficiaries moved inversely with the mean number of serious conditions diagnosed in 306 hospital referral regions; paradoxically, the greater the number of serious conditions diagnosed, the lower the fatality rate.²²

Several potential explanations exist for the case mix increase among intervention subjects. One possibility is that the latter scores represent a more valid estimate of comorbidity than at baseline because both groups had previously been underserved. This interpretation is supported by evidence from a variety of qualitative and quantitative sources. First, at least 82% met criteria to be considered underinsured relative to their ability to pay for care. Participants in 6 fo-

cus groups (n = 42) conducted during the first 2 years of our study reported forgoing or deferring care for complex and potentially serious conditions, such as kidney disease and gastrointestinal bleeding, and employing numerous strategies to minimize cost, such as saving up procedures until they had met deductibles and rationing medications.¹⁶ Once the DMIE program benefits removed the cost barriers, they reported that they were able to afford needed services. We examined the records of individuals with large changes in ACG weights and found baseline self-reported conditions, such as diabetes and a cancer history, for which they had no claims during the baseline year but which were treated after the intervention began.

The time-limited nature of the intervention may have also encouraged those individuals with pent-up need to consume high levels of service while the DMIE benefit was in place. Although all types of service consumption increased, we saw notable spikes in elective surgeries, such as joint replacements, expensive screening and diagnostic testing, and services not included or for which coverage was very limited under the basic state high-risk pool benefit. The latter include prescription coverage with cost sharing of 50% after deductible that decreased to a flat \$3 per drug copay under the DMIE, a major benefit for individuals heavily dependent on high-cost prescriptions for conditions such as cancer and autoimmune diseases.

■ **Table 5.** Within Group Percentage Changes in ACG Acuity Over a 2-Year Period

Highest Level of Acuity (ACG Actuarial Cell)	Intervention			Control		
	Baseline	Year 1	Year 2	Baseline	Year 1	Year 2
≤10 ADGs (5010-5070)	16.3	28.0	32.2	12.0	10.5	15.0
6-9 ADGs (4510-4940)	25.5	41.2	40.4	31.6	25.6	29.3
4-5 ADGs (4210-4430)	20.7	14.8	12.0	18.8	19.5	10.5
2-3 ADGs (3800-4100)	16.3	10.4	8.7	14.3	19.5	18.0
Acute (100-400, 1800-3700)	10.3	2.7	4.9	8.3	10.5	12.8
Chronic (800-900, 1200)	3.3	—	0.5	3.8	4.5	0.8
Psychiatric (1300-1500)	0.5	1.6	0.5	1.5	2.3	3.0
Preventive (1600)	0.5	—	—	1.5	2.3	1.5
Likely to recur (500-600)	1.1	—	—	—	0.8	—
Eye and dental (1100)	—	0.5	—	—	—	0.8
Nonusers (5100-5200)	5.4	0.5	0.5	8.3	4.5	8.3
Total	100.0	100.0	100.0	100.0	100.0	100.0

ACG indicates adjusted clinical group; ADG, aggregated diagnostic group.

In addition, all intervention group members received telephonic case management from registered nurses, who assisted subjects in identifying unmet needs and provided prior approval of insurance coverage. Although this surveillance was intended to help subjects optimize their health status while preventing wasteful utilization, it also inevitably raised awareness of unmet needs and stimulated a demand for services. Physicians who became aware of the availability of increased coverage also may have made more recommendations for interventions during this time-limited benefit period.

Because of this spike in utilization, some of which represented pent-up demand, the resulting case-mix scores may be inflated. However, many of the conditions for which participants sought treatment are chronic and represent continued clinical and actuarial risk, whether or not they continue to be incorporated into risk scores. This temporal aspect of risk adjustment highlights another limitation: case-mix scores normally encompass only 1 year of data; pre-existing conditions that have not been treated during the past year will not be detected, and underlying comorbidity may be vastly underestimated.

Study Limitations

The small sample size and high comorbidity risk of this study population, consisting of underinsured individuals with intensive need for healthcare services, limits the generalizability to large health plans with more generous health benefits and healthier, or more typically representative, health status. In addition, this study was conducted in only 1 small state; however, similar case mix increases were observed in

the Texas DMIE, which used the ACG with a previously uninsured group of 1400 individuals whose initial health status was similar to those in the Kansas DMIE study.²³ The findings seem most relevant to populations with high unmet health-care need who suddenly gain access to relatively comprehensive healthcare benefits.

CONCLUSIONS

The ACA proposes to use risk adjustment as a permanent mechanism for protecting insurance plans against adverse selection. The federal government will provide insurance exchanges with a risk-adjustment tool, such as the HCC, or allow a comparable substitute. But all of these tools share important vulnerabilities. Although intended to shield plans against adverse selection, these tools may inadvertently reward the inefficient at the expense of the efficient. At worst, they could encourage insurance industry “gaming” in order to maximize revenues and shareholder profits.²⁴

Claims-based tools may be particularly inaccurate when beneficiaries are chronically ill and underinsured. When many Americans gain coverage for the first time under the ACA, these factors may be an important issue in understanding risk scores. Many uninsured Americans are of lower socioeconomic status and in poor health.²⁵ Those who enroll in exchanges will still have cost-sharing barriers, particularly if they are above the 200% federal poverty threshold, where coinsurance is no longer subsidized. If they do not attain adequate access to services, their comorbidity risk may be underestimated by case mix–based risk scores. At this writing, how

the federal government will make use of individual risk scores is unsettled, and the ACA includes other interim financial safeguards in the form of reinsurance and risk corridors. However, policy makers, as well as others who contemplate the use of these tools for measuring comorbidity, need to be aware of this potential threat to their validity.

Author Affiliations: From Institute for Health and Disability Policy Studies, University of Kansas (JMM, JPH), Lawrence, KS.

Funding Source: This study was funded by the Kansas Health Policy Authority through a grant from the US Department of Health and Human Services, Centers for Medicare & Medicaid Services (No. 11-P-92389/7-01).

Author Disclosures: The authors (JMM, JPH) report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (JMM, JPH); acquisition of data (JMM, JPH); analysis and interpretation of data (JMM, JPH); drafting of the manuscript (JMM, JPH); critical revision of the manuscript for important intellectual content (JMM, JPH); statistical analysis (JMM); obtaining funding (JPH); administrative, technical, or logistic support (JPH); and supervision (JPH).

Address correspondence to: Jean P. Hall, PhD, MSW, JR Pearson Hall, Rm 517, 1122 W Campus Rd, Lawrence, KS 66045-3101. E-mail: jhall@ku.edu.

REFERENCES

- Ash A, Ellis R, Pope G, et al. Using diagnoses to describe populations and predict costs. *Health Care Finance Rev.* 2000;21(3):7-28.
- Weiner J, Starfield B, Steinwachs D, et al. Development and application of a population-oriented measure of ambulatory care case-mix. *Med Care.* 1991;29(5):452-472.
- Kronick R, Gilmer T, Dreyfus T, Lee L. Improving health-based payment for Medicaid beneficiaries. *Health Care Finance Rev.* 2000;21(3):29-64.
- Iezzoni L, ed. Risk adjustment for measuring health care outcomes. Chicago: Health Administration Press; 2003.
- Risk based capital guidelines: proposed rule. *Federal Register* 76:7 (11 Jan 2011) 1890-1922.
- Hall M. Risk adjustment under the affordable care act: a guide for federal and state regulators. May 2011. Commonwealth Fund Issue Brief, pub 1501, vol 7.
- Klabunde C, Warren J, Legler J. Assessing comorbidity using claims data: an overview. *Med Care.* 2002;40(8 suppl):IV-26-IV-35.
- Quan H, Parson G, Ghali W. Validity of information on comorbidity derived from ICD-9-CCM Administrative Data. *Med Care.* 2002;40(8):675-685.
- Iezzoni LI. Assessing quality using administrative data. *Ann Int Med.* 1997;127(8, pt 2):666-674.
- Choudry N, Shrank W. Four-dollar generics—increased accessibility, impaired quality assurance. *N Engl J Med.* 2010;363(20):1885-1887.
- Gimm G, Ireys H, Gilman G, et al. Impact of early intervention programs for working adults with potentially disabling conditions: evidence from the national DMIE evaluation. *J Voc Rehab.* 2010;34(2):67-69.
- Hall JP, Moore JM, Welch, G. Preventing disability among working participants in Kansas' high-risk insurance pool: Implications for health reform. *J Vocat Rehabil.* 2011;34(2):119-128.
- Hall JP, Moore, JM. Does high risk pool coverage meet the needs of a population at risk for disability? *Inquiry.* 2008;45(3):340-352.
- Collins S, Kriss J, Davis K, Doty M, Holmgren A. Squeezed: why rising exposure to health care costs threatens the health and financial well-being of American families. Publication No. 953. New York: The Commonwealth Fund. http://www.commonwealthfund.org/~/media/Files/Publications/Fund%20Report/2006/Sep/Squeezed%20%20Why%20Rising%20Exposure%20to%20Health%20Care%20Costs%20Threatens%20the%20Health%20and%20Financial%20Well%20Being%20of%20Collins_squeezedrisinghlcarecosts_953%20pdf.pdf. Published 2006. Accessed Dec. 5, 2011.
- Schoen C, Collins S, Kriss J, et al. How many are underinsured? trends among U.S. adults, 2003 and 2007. *Health Aff.* 2008;27(4):w298-w309.
- Hall JP, Carroll S, Moore JM. Health care behaviors and decision-making processes among enrollees in a state high-risk insurance pool: Focus group findings. *Am J Health Promot.* 2010;24(5):304-310.
- Ware J, Kosinski M, Keller S. A 12-item short-form health survey: construction of scales and preliminary tests of reliability and validity. *Med Care.* 1996;34(3):220-233.
- Centers for Disease Control and Prevention. National health interview survey on disability. http://www.cdc.gov/nchs/nhis/nhis_disability.htm. Updated June 17, 2009. Accessed Apr. 26, 2012.
- Weiner J, Abrams C, Millman A, eds. (2006). The Johns Hopkins ACG Case-Mix System: Technical User Guide, Version 8.0. Baltimore, MD: Johns Hopkins University, Bloomberg School of Public Health, 7-2.
- Ware J, Kosinski M. SF-36 *Physical and Mental Health Summary Scales: A Manual for Users of Version 1*. 2nd ed. Lincoln, RI: Quality Metric, Incorporated; 2001.
- Song Y, Skinner J, Bynum J, Sutherland J, Wennberg J, Fisher E. Regional Variations in Diagnostic Practices. *N Engl J Med.* 2010; 363(1):45-53.
- Welch H, Sharp S, Gottlieb D, Skinner J, Wennberg J. Geographic variation in diagnosis frequency and risk of death among Medicare beneficiaries. *JAMA.* 2012;305(11):1113-1118.
- Bohman T, Wallisch L, Spence R, et al. Examining ACG health measures as outcomes in longitudinal research within a public health system. http://www.acg.jhsph.edu/acgdocuments/Tuesday_Boardroom_Bohman_AnalyticIssues.pdf. Presented May 11, 2010. Accessed April 25, 2012.
- Weiner J, Trish E, Abrams C, Lemke K. Adjusting for risk selection in state health insurance exchanges will be critically important and feasible, but not easy. *Health Aff.* 2012;31(2):306-315.
- Kaiser Family Foundation, The Henry J. A profile of health insurance exchange enrollees. <http://www.kff.org/healthreform/8147.cfm>. Published 2011. Accessed Dec. 5, 2011. ■