

Patient Complexity in Quality Comparisons for Glycemic Control: an Observational Study

Monika Safford, MD^{1,2}; Michael Brimacombe, PhD^{1,3}; Quanwu Zhang, PhD¹; Mangala Rajan, MBA¹; Minge Xie, PhD^{1,4}; Wesley Thompson, PhD^{1,4}; John Kolossa, PhD^{1,4}; Miriam Maney¹; Leonard Pogach, MD, MBA¹

¹VA New Jersey Healthcare System, East Orange, NJ

²Deep South Center on Effectiveness at Birmingham VA Medical Center and University of Alabama at Birmingham, Birmingham, AL

³University of Medicine and Dentistry of New Jersey-New Jersey Medical School, Newark, NJ

⁴Rutgers University, Piscataway, NJ

Corresponding author:

Monika Safford, MD

MT643 1717 11th Av. S.

Birmingham, AL 35205

(205)934-6883 Fax: (205)934-7959

Presented at the American Diabetes Association, San Francisco, CA June 17, 2002.

Supported by a generous grant from the American Diabetes Association (Dr. Safford, PI) and VA Health Services Research & Development Investigator Initiated Research award 00-072-1 (Dr. Pogach, PI)

2,296 words 3 Tables 3 Figures

Running title: Adjusting A1c for patient complexity

Author emails:

Monika Safford, MD	msafford@uab.edu
Michael Brimacombe, PhD	brimacmb@umdnj.edu
Quanwu Zhang, PhD	quanwu.zhang@aventis.com
Mangala Rajan, MBA	rajan@njneuromed.org
Minge Xie, PhD	mxie@stat.rutgers.edu
Wesley Thompson, PhD	wesleyt@pitt.edu
John Kolassa, PhD	kolassa@stat.rutgers.edu
Miriam Maney	maney@njneuromed.org
Leonard Pogach, MD, MBA	leonard.pogach@med.va.gov

Abstract

Background: Patient complexity is not incorporated into quality of care comparisons for glycemic control. We developed a method to adjust hemoglobin A1c levels for patient characteristics that reflect complexity, and examined the effect of using adjusted A1c values on quality comparisons.

Methods: This cross-sectional observational study used 1999 national VA pharmacy, inpatient and outpatient utilization and laboratory data on diabetic veterans. We adjusted individual A1c levels for available domains of complexity: age, social support (marital status), comorbid illnesses, and severity of disease (insulin use). We used adjusted A1c values to generate VA Medical Center level performance measures, and compared Medical Center ranks using adjusted vs. unadjusted A1c levels across several thresholds of A1c (8.0%, 8.5%, 9.0% and 9.5%).

Results: The adjustment model had $R^2 = 8.3\%$ with stable parameter estimates on thirty random 50% resamples. Adjustment for patient complexity resulted in the greatest rank differences in the best and worst performing deciles, with similar patterns across all tested thresholds.

Conclusions: Adjustment for complexity resulted in large differences in identified best and worst performers at all tested thresholds. Current performance measures of glycemic control may not be reliably identifying quality problems, and tying reimbursements to such measures may compromise the care of complex patients.

Key Words: diabetes, quality of care, complexity, hemoglobin A1c

Background

Patient complexity has recently been raised as an important issue in patient care and quality assessment [1-4], and in diabetes, complexity can arise from multiple sources. Because many type 2 diabetes patient are older, they frequently have multiple medical problems, and complexity arising from multiple comorbid conditions has been discussed by several authors [1-4]. Diabetes can also be particularly challenging to manage when it is advanced, calling for insulin self-administration and frequent self-testing. Accomplishing the considerable self management demands of diabetes [5] can be especially difficult for patients who lack social support [6-8]. In the VA, married diabetic men have better glycemic control than unmarried diabetic men [9]. All of these sources of complexity call for clinical trade-off decisions, possibly deviating from “ideal care” for any one disease state. When the Institute of Medicine recommended that the ideal health care system deliver care that is “driven by shared decision making and based on continuous, healing relationships” [10], it acknowledged that for some patients, especially those that are complex, “good care” will not necessarily lead to ideal performance measures.

None of these aspects of patient complexity are currently reflected in performance measures widely used for public reporting, and, more recently, to “pay for performance” (P4P) [11-13]. While public reporting of the quality of healthcare has had measurable effects on improving population health [14-16], currently implemented performance measures focus on one disease at a time, raising concern about how well they assess the overall quality of care of complex patients [1-4]. For diabetes, the public

accountability measure has been an assessment of poor glycemic control as reflected by hemoglobin A1c >9.0%, with plans to set this threshold at lower levels [17]. This performance measure is based on well-established evidence that glycemic control is associated with diabetes outcomes [18]. However, it is unclear how patient complexity influences the assessment of quality of care provided by health care plans.

We studied how incorporating several readily available characteristics that reflect patient complexity (age, comorbidity, severity of diabetes as reflected in insulin use, and social support as reflected by marital status) affected assessments of quality of care. We conducted this study in the nation's largest integrated health system, the Veterans Health Administration (VHA).

Methods

Data

We used data from the VHA's Healthcare Data Analysis Information Group in Milwaukee, Wisconsin for pharmacy data and A1c values, and the National Patient Clinical Dataset from the Veterans Integrated Service Network Support Center in Austin, Texas for inpatient and outpatient administrative utilization data with associated International Classifications of Diseases, 9th edition (ICD-9) codes and demographic information. To identify diabetes, we selected veterans who used the VHA in 1998, were alive at the end of 1999, and in 1999 received either a diabetes medication or had an ICD-9 code 250.xx (diabetes) associated with more than one outpatient encounter or any inpatient encounter at one of the VHA's 145 Medical Centers [19].

When multiple values for A1c were present for an individual for 1999, we used the last available value. Values that fell above the physiologic range were excluded (>18.0%).

Several VA Medical Centers could not be included because of A1c lab assay methodology. Precision and bias problems with A1c laboratory assays to measure A1c levels have been reported [20-22] and a national effort to standardize A1c test methods is underway. Since standardization of A1c methodology was not mandated in the VHA until late in 1999, we contacted the laboratory director at each Medical Center to determine which A1c lab method was utilized at their Medical Center that year. Only the 66 Medical Centers using exclusively National Glycohemoglobin Standardization Project-certified methods were included in this study [23].

Variables used to reflect patient complexity

Patient complexity can arise along any of the axes of the major determinants of health: socio-economic, cultural, biological, environmental, and behavioral [24]. Not all of these influences can be readily assessed using administrative data. Because our goal was to use existing data, we examined variables that were available in the VA databases. They included age, married status, comorbid illnesses and severity of diabetes approximated by insulin treatment. In type 2 diabetes, insulin treatment signals failure of oral therapies and reflects more advanced disease. Patients are often reticent to initiate treatment requiring needles and more intense monitoring, therefore insulin treatment in type 2 patients is an approximate indicator of disease severity [18]. To represent comorbid illnesses, we used a method developed among veterans that sums

the presence of any of 30 common chronic illnesses into a single unweighted score [25]. Both inpatient and outpatient utilization data and associated ICD-9 codes were used to construct the comorbidity score.

Modeling

We used linear regression models to complexity-adjust individual A1c levels as a continuous measure. Covariates included age, marital status, insulin treatment, and comorbidity score. We also considered interactions between age and the comorbidity score, and between age and diabetes treatment. We retained only variables that were significant at the $p < .05$ level in the final model.

We evaluated the model's performance by examining R^2 and by dividing the adjusted A1c values into deciles, then examining the proportion of the unadjusted above-threshold values in each decile [26]. We evaluated model stability by drawing 30 random 50% subsamples and examining their ranges of the regression coefficients and R^2 .

Profiling

To evaluate the effect of complexity-adjustment on profiling (rank order), we proceeded in 3 steps (see Figure 1). In Step 1, we created "observed," or unadjusted ranks. We first created proportions for each VAMC consisting of the number of individuals with observed A1c at or above threshold divided by the total number of diabetes patients at that VAMC. We then ranked VAMC on these proportions.

In Step 2, we created “observed-to-expected,” or adjusted, ranks. To do this, we first determined the entire national study population’s (all 66 VA Medical Centers) proportion of observed uncontrolled patients (the national observed uncontrolled proportion). Then, we flagged individuals with adjusted A1c’s at or above the corresponding national observed uncontrolled proportion for each threshold (the “expected” uncontrolled). For example, 15.6% of patients at the 66 Medical Centers overall had unadjusted A1c>9.5%; we therefore identified all sample members with the 15.6% highest adjusted A1c values. Next, we counted the number of individuals at each VAMC who were flagged as being in the national observed uncontrolled proportion. We then used this number of “expected” uncontrolled to create the “observed-to-expected” ratio for each VAMC. Last, we ranked VAMC on these “observed-to-expected” ratios.

In Step 3, we compared the results with ranks obtained in Step 2 (adjusted ranks) with those obtained in Step 1 (unadjusted ranks). We repeated this process for each of the tested thresholds of A1c: 8.0%, 8.5%, 9.0%, and 9.5%.

All statistical analyses were performed using STATA (Version 7.0, Stata Corporation, 2001) and SAS (Version 9.0, Cary, NC). The VA New Jersey Healthcare System institutional review board approved the study.

Results

The patients in the study sample (n=118,167) were similar to the overall VHA diabetes population (Table 1), with a mean age of 64 years (SD 11) and 63% married. Patients in the study sample had on average 1.7 more comorbid medical conditions than the population of veterans with diabetes. Thirty-nine percent were on insulin, either

alone or in combination with oral agents. The 66 sample VA Medical Centers cared for a mean number of 1790 patients with diabetes (range 328-5192) and 15.6% of the study sample had A1c >9.5% (range 4.9-25.2%), 21.7% >9.0% (6.7-32.1%), 29.6% >8.5% (11.9-42.3%), and 39.7% >8.0% (21.2-54.5%).

Medical Center ranks differed modestly simply by using a different A1c thresholds, without any adjustment. Compared with ranks obtained using the 9.0% threshold, 51% of ranks obtained using 8.0% were within 5 ranks. For the best quartile of performance, 76% of ranks for the 8.0% threshold were within 5 ranks of those obtained using 9.0%.

Modeling results

All variables and the age * comorbidity score interaction contributed significantly to variation in A1c and were retained in the final model (Table 2). This model's R² was 8.3%. The resampled coefficient means were very close to the original model's with narrow ranges (Table 2), reflecting model stability.

Decile of risk tables [26] indicated that each decile of adjusted A1c had successively more above-threshold observed values, as expected (Figure 2). Trends were similar for each threshold. The considerable change in the proportion of unadjusted above-threshold values across the deciles indicated that assessments based on unadjusted values were quite different from those based on adjusted values.

Effect of complexity-adjusted A1c on VA Medical Center profiling

The effects of adjusting for complexity on profiling were substantial. Figure 3 demonstrates that the greatest changes in rank occurred in the extreme deciles that are the focus of quality assessment. These changes were similar across all the thresholds tested. For Medical Centers in the best decile of performance, the average change in rank was 25, with similar magnitude of change for Medical Centers within the worst decile of performance. Table 3 depicts the actual rank changes experienced by the top and bottom 10 performers. Between 0 and 2 Medical Centers remained in the top 10 after complexity-adjustment, and 1 to 2 Medical Centers remained in the bottom 10, depending on the threshold. Remarkably, 2 to 3 Medical Centers in the bottom 10 became top 10 performers with complexity-adjustment, regardless of the threshold chosen.

Discussion

Using A1c values adjusted for only a few domains of patient complexity caused substantial differences in which VA Medical Centers were identified as best and worst performers. Depending on the threshold, 20-30% of the 10 best performers among these 66 Medical Centers using adjusted values would have been identified as among the 10 worst performers without adjustment. Our findings suggest that using unadjusted A1c in the glycemic control performance measure may not be reliably assessing quality of care. This finding calls into questions whether currently used methodology should be tied to reimbursement [12]; one of the indicators currently included in the Bridges to

Excellence diabetes program for physicians and providers includes the proportion of people with diabetes who had poorly controlled A1c levels at last measure.

Adjustment for complexity had a similar impact across all tested thresholds of A1c. Although our study focuses on thresholds in place at the time of the study (A1c>9.0%), our findings have implications for the recent decision to lower the threshold for glycemic control by National Committee on Quality Assurance, since there are no plans to include adjustment of A1c levels [17].

We were unable to include important domains of patient complexity, such health behaviors, cultural influences or socioeconomic influences in the adjustment model. Even the variables that we did include did not completely assess complexity per se. While the sheer number of medical conditions is important, the severity of the conditions is not reflected in ICD-9 codes. Some quality assessments currently rely on patient surveys, and it is likely that surveys could better capture important domains of not only patient complexity but also preferences. Surveys are not without their drawbacks, including biases and cost [27]. It is clear that thoughtful approaches to capturing the full picture of patient complexity are needed.

Nevertheless, our study's findings of large differences in identified best and worst performing Medical Centers underscores the urgency in incorporating complexity, even in rudimentary form, into quality comparisons. Because of the inevitable trade-off decisions required in the care of complex patients, their care may not be assessed as "guideline concordant." Tying guideline concordant care with reimbursement creates a tension for the treating physician: should s/he "treat for performance" or treat the patient? Both may not be possible; P4P could potentially pose a threat to the overall

quality of care received by complex patients, as perverse incentives encourage the clinician to spend less time engaging the patient in eliciting preferences and developing congruence on a tailored treatment plan.

The rudimentary variables we did include in our adjustment model are widely available, which could be abstracted from medical records during annual data collection for performance measures. The model demonstrated explanatory value in a range typical for case-mix adjustment models [26], the deciles of adjusted values had the expected increasing contribution of above-threshold unadjusted values, and the parameter estimates were stable with repeated sampling. Further, the simplicity of our approach makes it feasible in various settings.

Our study was conducted in the VA, which includes largely older men who may be more debilitated than the general population. Our findings should be examined in other populations.

Conclusions

Adjusting A1c levels for readily available characteristics that reflect patient complexity resulted in large differences in identifying best or worst performers, most pronounced at the extremes of performance that are the focus of quality assessment. These findings were similar across all tested thresholds of A1c, suggesting that patient complexity is an important influence at all levels of glycemic control. Current practices for assessing glycemic control as a quality of care indicator may be identifying differences in the populations health systems serve as much as differences in the quality of care they provide. Tying such measures to reimbursement may pose a threat

to the care of complex patients, and methods to adjust performance measures for complexity should be implemented.

Declaration of Competing Interests

The authors declare that they have no competing interests

Authors' Contributions

MS conceived of the paper, interpreted data, drafted the manuscript and procured funding. MB provided conception and design input, conducted the analyses and participated in providing critical revisions to the draft. QZ provided conception and design input, as well as critical revisions to the draft. MR provided conception and design input, and acquired data. MX, WT, and JK provided conception and design input, and provided critical revisions to the draft. MM conducted analyses and provided critical revisions to the draft. LP provided input into conception of the paper, interpreted data, critical revisions to the manuscript and procured funding.

Acknowledgement

We thank Arlene Ash for her helpful suggestions.

Competing interests

None of the authors have any competing interests, financial or otherwise.

References

1. Boyd CM, Darer J, Boult C, Fried LP, Boult L, Wu AW: **Clinical practice guidelines and quality of care for older patients with multiple comorbid diseases: implications for pay for performance.** *JAMA* 2005, **294**(6):716-724.
2. Brown AF, Mangione CM, Saliba D, Sarkisian CA: **Guidelines for improving the care of the older person with diabetes mellitus.** *J Am Geriatr Soc* 2003, **51**(5 Suppl Guidelines):S265-280.
3. Durso SC: **Using clinical guidelines designed for older adults with diabetes mellitus and complex health status.** *JAMA* 2006, **295**(16):1935-1940.
4. Tinetti ME, Bogardus ST, Jr., Agostini JV: **Potential pitfalls of disease-specific guidelines for patients with multiple conditions.** *N Engl J Med* 2004, **351**(27):2870-2874.
5. Russell LB, Suh DC, Safford MA: **Time requirements for diabetes self-management: too much for many?** *J Fam Pract* 2005, **54**(1):52-56.
6. Anderson BJ, Cullen K, McKay S: **Quality of life, family behavior, and health outcomes in children with type 2 diabetes.** *Pediatr Ann* 2005, **34**(9):722-729.
7. Tomaka J, Thompson S, Palacios R: **The relation of social isolation, loneliness, and social support to disease outcomes among the elderly.** *J Aging Health* 2006, **18**(3):359-384.
8. van Dam HA, van der Horst FG, Knoop L, Ryckman RM, Crebolder HF, van den Borne BH: **Social support in diabetes: a systematic review of controlled intervention studies.** *Patient Educ Couns* 2005, **59**(1):1-12.
9. Zhang Q, Safford M, Ottenweller J, Hawley G, Repke D, Burgess JF, Jr., Dhar S, Cheng H, Naito H, Pogach LM: **Performance status of health care facilities changes with risk adjustment of HbA1c.[comment].** *Diabetes Care* 2000, **23**(7):919-927.
10. Institute of Medicine, Committee on Quality Health Care in America: **Crossing the quality chasm.** Washington, DC: National Academy Press; 2001.
11. **Bridges to Excellence for Physicians & Providers, NCQA Diabetes Care Performance Assessment Program**
[\[http://www.bridgestoexcellence.org/bte/physician/diabetesassess.htm\]](http://www.bridgestoexcellence.org/bte/physician/diabetesassess.htm)
12. **Fact sheet: Medicare "pay for performance (P4P)" initiatives, January 31, 2005** [\[http://www.cms.hhs.gov/media/press/release.asp?Counter=1343\]](http://www.cms.hhs.gov/media/press/release.asp?Counter=1343)
13. **Kaisernetwork.org Daily Health Policy Report. Medicare: Rep. Johnson calls for pay-for-performance Medicare reimbursement system for physicians, March 16, 2005**
[\[http://www.kaisernetwork.org/daily_reports/re_index.cfm?hint=3&DR_ID=28742\]](http://www.kaisernetwork.org/daily_reports/re_index.cfm?hint=3&DR_ID=28742)
14. Hannan EL, Wu C, Walford G, King SB, 3rd, Holmes DR, Jr., Ambrose JA, Sharma S, Katz S, Clark LT, Jones RH: **Volume-outcome relationships for percutaneous coronary interventions in the stent era.** *Circulation* 2005, **112**(8):1171-1179.
15. Jha AK, Perlin JB, Kizer KW, Dudley RA: **Effect of the transformation of the Veterans Affairs Health Care System on the quality of care.** *N Engl J Med* 2003, **348**(22):2218-2227.

16. Mukamel DB, Weimer DL, Mushlin AI: **Referrals to high-quality cardiac surgeons: patients' race and characteristics of their physicians.** *Health Serv Res* 2006, **41**(4 Pt 1):1276-1295.
17. **Quality Update for April 9, 2004. New NCQA Quality Diabetes Measures Endorsed** [http://www.ahqa.org/pub/connections/162_696_4750.cfm#newn]
18. Turner RC, Cull CA, Frighi V, Holman RR: **Glycemic control with diet, sulfonylurea, metformin, or insulin in patients with type 2 diabetes mellitus: progressive requirement for multiple therapies (UKPDS 49).** UK Prospective Diabetes Study (UKPDS) Group. *JAMA* 1999, **281**(21):2005-2012.
19. Hebert PL, Geiss LS, Tierney EF, Engelgau MM, Yawn BP, McBean AM: **Identifying persons with diabetes using Medicare claims data.** *Am J Med Qual* 1999, **14**(6):270-277.
20. Little RR, Wiedmeyer HM, England JD, Naito HK, Goldstein DE: **Interlaboratory comparison of glycohemoglobin results: College of American Pathologists Survey data.** *Clin Chem* 1991, **37**(10 Pt 1):1725-1729.
21. McKenzie PT, Sugarman JR: **Implications for clinical performance measurement of interlaboratory variability in methods for glycosylated hemoglobin testing.** *Am J Med Qual* 2000, **15**(2):62-64.
22. Takahashi Y, Noda M, Tsugane S, Kimura S, Akanuma Y, Kuzuya T, Ohashi Y, Kadowaki T: **Importance of standardization of hemoglobin A1c in the analysis of factors that predict hemoglobin A1c levels in non-diabetic residents of three distinct areas of Japan.** *Diabetes Res Clin Pract* 2001, **53**(2):91-97.
23. **NGSP** [<http://www.missouri.edu/~diabetes/ngsp/index.html>]
24. Tarlov AR: **Public policy frameworks for improving population health.** *Ann N Y Acad Sci* 1999, **896**:281-293.
25. Selim AJ, Fincke G, Ren XS, Lee A, Rogers WH, Miller DR, Skinner KM, Linzer M, Kazis LE: **Comorbidity assessments based on patient report: results from the Veterans Health Study.** *J Ambul Care Manage* 2004, **27**(3):281-295.
26. Schwartz M, Ash A: **Evaluating the performance of risk-adjustment methods: continuous outcomes.** In: *Risk Adjustment for Measuring Healthcare Outcomes*. Edited by Iezzoni LI, 2nd edn. Chicago, IL: Health Administration Press; 1997: 391-426.
27. **Polls face growing resistance, but still representative survey experiment shows** [<http://people-press.org/reports/display.php3?ReportID=211>]

Table 1. Patient characteristics* of all VHA patients with diabetes and the study sample.

Patient characteristic	Veterans with diabetes	Study sample
N	552,128	118,167
Age in years, mean \pm SD	64.2 \pm 11.1	64.1 \pm 11.1
Married, %	61.9	62.7
Diabetes treatment, %		
No VHA meds	3.3	3.2
Oral agents only	57.7	57.9
Insulin only	21.9	22.7
Insulin + oral agent(s)	17.2	16.2
Comorbidity score \pm SD	3.4 \pm 2.6	5.1 \pm 3.0

*p-values not shown due to large sample and small differences reaching statistical significance. 98% of the sample was male.

Table 2. Case-mix adjustment model for A1c with model coefficients and coefficients of thirty random 50% subsamples (with resampling).

Variable	Model Coefficients \pm SE	Mean of Resampled Coefficients [Range]
Age group (years)		
<55 vs. >75	.92 \pm .03	.92 [.84, .99]
55 – 65 vs. >75	.49 \pm .03	.49 [.39, .56]
65 – 75 vs. >75	.21 \pm .03	.20 [.14, .26]
Married vs. not	.07 \pm .01	.07 [.05, .10]
Insulin vs. no insulin	-.75 \pm .01	-.75 [-.78, -.72]
Comorbidity score	-.02 \pm .00	-.03 [-.04, -.02]
Age group * Comorbidity score		
<55 vs. >75	-.05 \pm .01	-.05 [-.07, -.04]
55 – 65 vs. >75	-.02 \pm .01	-.02 [-.03, .00]
65 – 75 vs. >75	-.01 \pm .01	-.01 [-.02, -.00]
Intercept	8.16 \pm .02	8.16 [8.10, 8.23]
Model R ²	.083	.084 [.078, .089]

Table 3. Rank changes* with adjustment among the top 10 and bottom 10 performers among 66 VA Medical Centers

	Adjusted rank			
	8.0%	8.5%	9.0%	9.5%
Top 10 performers, unadjusted rank				
1	30	7	3	29
2	11	33	31	20
3	40	40	30	11
4	56	34	48	3
5	36	56	9	56
6	29	17	35	38
7	39	23	28	46
8	51	42	50	33
9	22	41	32	37
10	37	48	43	28
Number among top 10 performers, unadjusted, who would be ranked as top 10 performers with adjustment				
	0	1	2	1
Bottom 10 performers, unadjusted rank				
57	50	58	5	34
58	4	50	12	21
59	3	60	57	60
60	58	51	59	10
61	52	53	53	5
62	57	3	56	55
63	8	55	4	2
64	20	16	23	26
65	38	6	44	27
66	54	49	6	44
Number among bottom 10 performers, unadjusted, who would be ranked as <i>top</i> 10 performers with adjustment				
	2	2	3	3

*Shaded cells indicate adjusted ranks in the top 10.

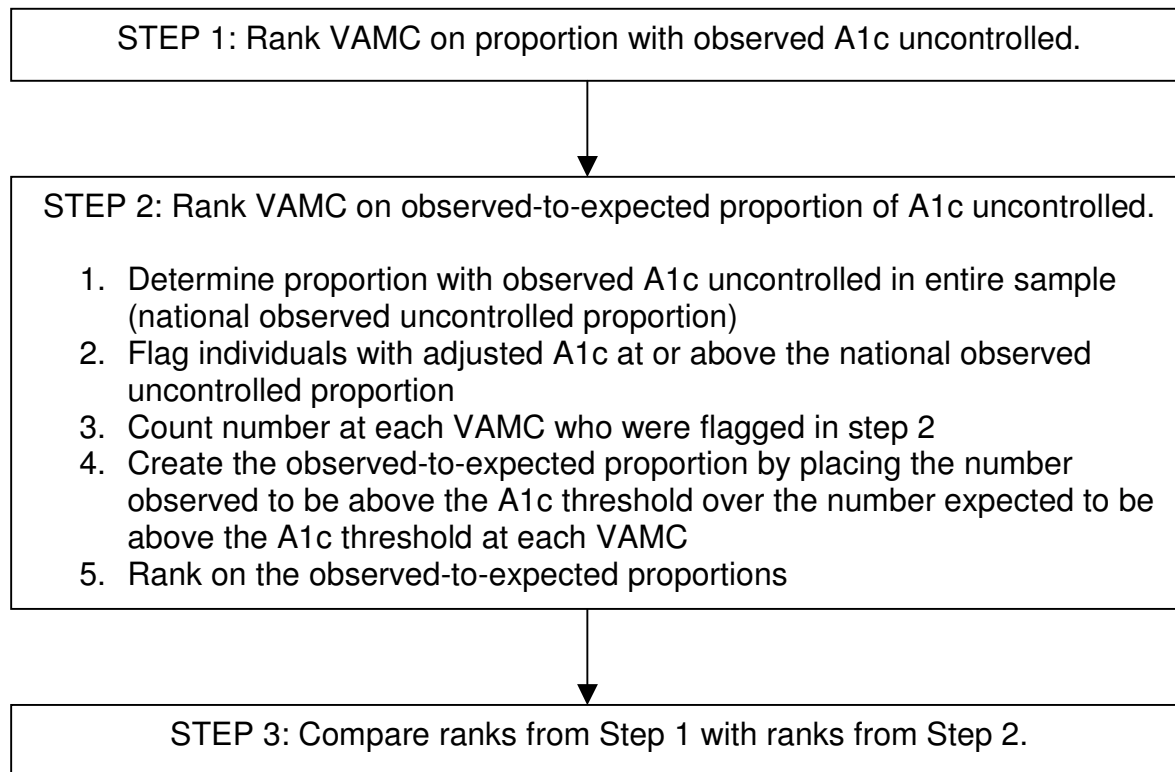


Figure 1. Steps used to examine results of using unadjusted vs. adjusted A1c to rank VA Medical Centers on glycemic control.

Percent above-threshold unadjusted A1c in each decile of adjusted A1c

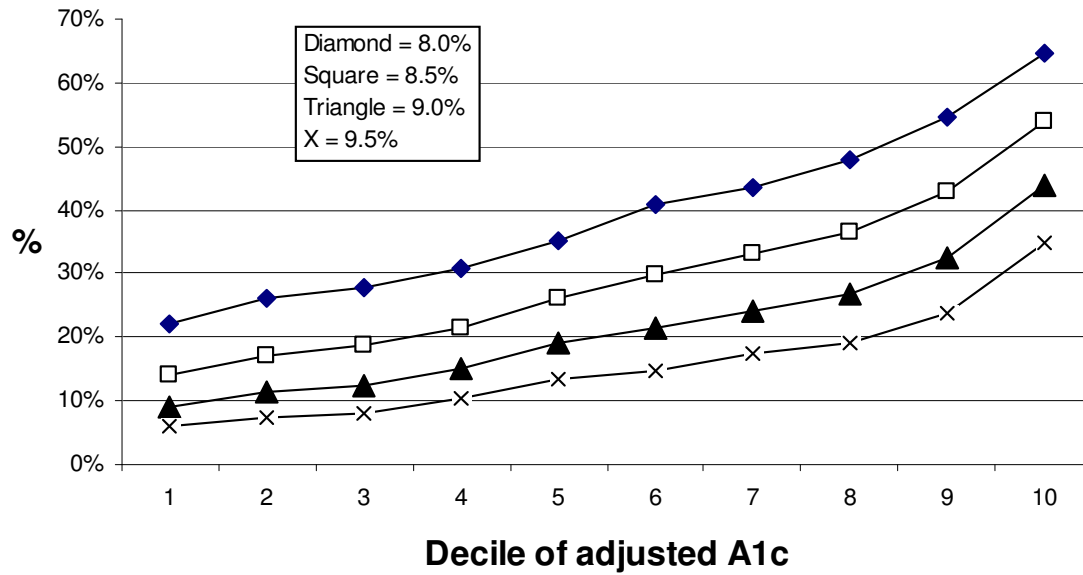


Figure 2. Deciles of adjusted A1c with contribution of above-threshold unadjusted A1c.

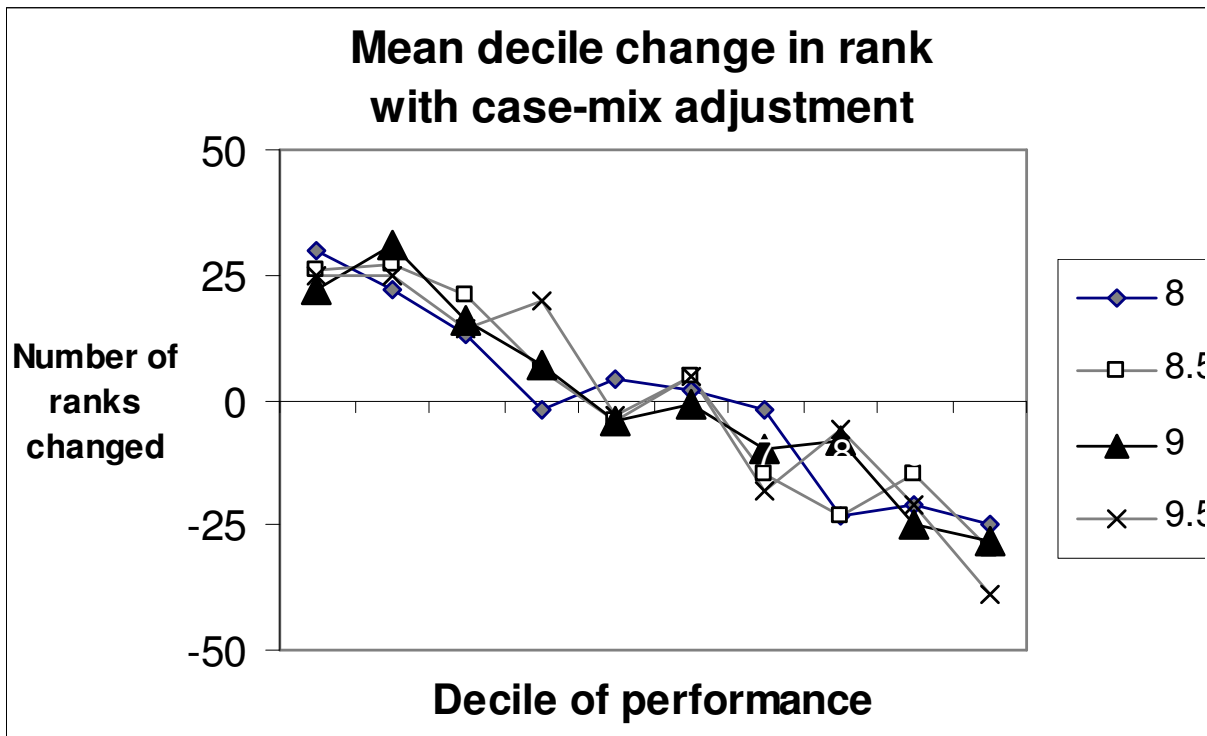


Figure 3. Mean decile change in rank with case-mix adjustment for 66 VA Medical Centers.

Legend to Figure 3.

Decile 1 includes the “top” performers among these 66 VA Medical Centers. For each threshold of A1c, VA Medical Centers in Decile 1 experienced an average change in rank of 25.