Hidden Complexities in Assessment of Glycemic Outcomes

Are quality rankings aligned with treatment?

LEONARD M. POGACH, MD, MBA[1,2]
MANGALA RAJAN, MBA[1]
MIRIAM MANEY, MA[1]
CHIN-LIN TSENG, DRPH[1,2]
DAVID C. ARON, MD, MS[3,4]

OBJECTIVE — To evaluate facility rankings in achieving <7% A1C levels based on the complexity of glycemic treatment regimens using threshold and continuous measures.

RESEARCH DESIGN AND METHODS — We conducted a retrospective administrative data analysis of Veterans Health Administration Medical Centers in 2003–2004. Eligible patients were identified using National Committee for Quality Assurance (NCQA) measure specifications. A complex glycemic regimen (CGR) was defined as receipt of insulin or three oral agents. Facilities were ranked using five ordinal categories based up both z score distribution and statistical significance ($P < 0.05$). Rankings using the NCQA definition were compared with a subset receiving CGRs using both a <7% threshold and a continuous measure awarding proportional credit for values between 7.9 and <7.0%. Ranking correlation was assessed using the Spearman correlation coefficient.

RESULTS — A total of 203,302 patients (mean age 55.2 years) were identified from 127 facilities (range 480–5,411, mean 1,601); 26.7% (17.9–35.2%) were receiving CGRs, including 22.0% receiving insulin. Mean A1C and percent achieving A1C <7% were 7.48 and 48% overall and 8.32 and 24.8% for those receiving CGRs using the threshold measure; proportion achieved was 60.1 and 37.2%, respectively, using the continuous measure. Rank correlation between the overall and CGR subset was 0.61; 8 of 24 of the highest or lowest ranked facilities changed to nonsignificance status; an additional five sites changed rankings.

CONCLUSIONS — Facility rankings in achieving the NCQA <7% measure as specified differ markedly from rankings using the CGR subset. Measurement for public reporting or payment should stratify rankings by CGR. A continuous measure may better align incentives with treatment intensity.

The National Committee for Quality Assurance (NCQA)-Healthcare Employer Information Data Set (HEDIS) measure for good (<7% A1C) glycemic control was revised in 2009 to apply only to individuals <65 years without cardiovascular diseases, end-stage diabetes complications, or dementia (1). Although most patients will require the use of insulin to achieve tight control within 10 years of diagnosis (2), even short-term attainment of an A1C level <7% may be difficult for such individuals (3,4). In a representative sample of the U.S. population, receipt of insulin resulted in a 14% decrease in the number of individuals attaining a <7% threshold (5). Because the frequency and risk of hypoglycemia may be a concern (6), multiple guidelines recommend that glycemic control targets should be individualized for such individuals, especially in the presence of comorbid medical and mental health conditions (7). Furthermore, individuals with diabetes may choose not to initiate insulin for suboptimal glycemic control even when recommended (8). Thus, any performance measure, especially one used for public reporting or provider payment, should take into account case mix differences among populations. Although there are no validated risk adjustment models for glycemic control, one potential indicator of case mix is the type and complexity of the medication regimen used (5).

The Veterans Health Administration (VHA) is the nation's largest integrated health care system and has national data repositories of individual administrative, pharmacy, and laboratory records (9). Our study objective was to compare changes in facility-level rankings based on the NCQA-HEDIS measure, as currently specified, with rankings based on a subset of patients receiving either insulin or three oral agents. We hypothesized that the identification of "best" and "worse" performing facilities would differ between these groups. Additional objectives were to compare rankings using the threshold measure with those using a continuous measure that awarded incremental credit for improvement between 7.9 and <7.0% A1C (10), and to evaluate the impact of excluding additional serious medical and mental illnesses.

RESEARCH DESIGN AND METHODS

Data sources and patient identification

Inpatient and outpatient utilization data, ICD-9-CM codes, and diagnostic codes were obtained from the National Patient Clinical Dataset (Austin, TX), laboratory data was obtained from the Decision Support System, and medication data were from the Pharmacy Benefits Management Program (11). To approximate NCQA criteria for continuous enrollment, patients were eligible if they received care within the VHA in fiscal year 2003 (1 October

2002 to 30 September 2003), received VHA care in FY2004, and had a diagnosis of diabetes. The latter was defined as two or more diabetes ICD-9 codes (250, 357.2, 362.0, and 366.41) associated with clinical outpatient care on separate calendar days or any diabetes-specific medication prescription (insulin, sulfonylureas, biguanides, α-glucosidase inhibitors, meglitinides, and thiazolidinediones). Individuals for whom death was ascertained before 30 September 2004 were excluded. We eliminated veterans without an A1C test in VHA databases because results done outside the VHA cannot be differentiated from nonperformance of the test.

Patients were defined as having a complex glycemic regimen (CGR) if they were receiving either insulin or three oral agents. Such patients would be considered to have received maximal oral agent therapy, especially if they were receiving thiazolidinediones. During this period, the VHA Pharmacy Benefits Management (12) classified thiazolidinediones as third-line therapy and facilities monitored their usage.

Denominator exclusion criteria

The denominator for the primary analyses was based on current NCQA-HEDIS criteria that exclude individuals with cardiovascular conditions, end-stage complications of diabetes (blindness, amputations, and end-stage renal disease), and dementia based on one ICD9-CM or Current Procedural Terminology (CPT) code in either the calendar year or preceding year (1). To evaluate the impact of other chronic illnesses on facility rankings, we similarly identified and excluded patients in the NCQA denominator for conditions in the following categories, as previously described (13): decreased life expectancy (cancers [except basal and squamous skin cancers] and end-stage hepatic disease); advanced (but not end-stage) complications of diabetes (advanced retinopathy [proliferative disease, prior laser treatment, retinal detachment, macular edema, and vitreous hemorrhage] without blindness); chronic renal disease; gastroparesis; potentially disabling neurological diseases (Parkinson's disease and spinal cord injury); major mental health conditions (major depression, schizophrenia, and bipolar disorder); and substance abuse. Patients with these conditions were considered to have decreased benefit and/or increased risk from intensive treatment (7).

Analysis

We evaluated the distributions of glycemic control for all patients in the NCQA denominator, as well as for patients stratified by CGR and non-CGR status (using both 0.1% A1C increments and ordinal intervals). At the facility level, we determined the proportion of patients who achieved <7% A1C based upon the last A1C test in FY04 within the specified denominators. To increase the reliability of facility rankings, we specified that a facility must have at least 100 patients receiving CGRs (14).

In addition to the use of the NCQA measure, we also evaluated the use of a continuous measure that would provide partial credit for A1C values between 7.9 and 7.0% and full (100%) credit for values <7.0% (10). In contrast, a threshold measure (such as <7% A1C) awards 100% credit only for values <7% and no credit for any values ≥7%. In brief, because the lifetime absolute risk reduction of microvascular complications resulting from lower A1C is linear within this range (as opposed to a log-linear absolute risk reduction when starting at higher A1C levels), arithmetic interpolation was used to equally weight each 0.1% decrease in A1C.

Facility rankings were determined by assigning facilities into five ordinal categories based on both z score distribution (≤10% [five stars or best], 10–33% [four stars], 34–66% [three stars], 67–90% [two stars], and ≥90% [one star or worst]) and statistical significance ($P <$ 0.05) (15). Statistical significance was calculated using the interval as follows:

$$(\text{Facility rate} - \text{Overall mean rate})$$
$$\pm 1.96\sqrt{(\text{SE facility})^2}$$
$$+ (\text{SE overall})^2$$

Significance was assumed if the interval did not include zero. Differences in ranking were assessed using the Spearman correlation coefficient and movement of facilities between the categories. The study was approved by the institutional review boards of the East Orange and Cleveland VA Medical Centers.

RESULTS — We identified 378,112 patients with diabetes who were <65 years of age at the onset of FY2004 (supplementary Fig. 1, available in an online appendix at http://care.diabetesjournals.org/cgi/content/full/dc09-1665). We excluded 63,753 patients because A1C results were not available in VHA databases. An additional 108,901 patients were excluded based on the NCQA technical measure specifications. Finally, 2,156 patients were excluded from six sites with <100 patients receiving CGRs. The final study sample was 203,302 patients from 127 facilities, 54,351 (26.7%) of whom received a CGR, including 22.0% receiving insulin (Table 1). Of patients receiving three oral agents, 93.7% were taking a thiazolidinedione. The facility range for CGRs was 17.9–35.2% (not shown). The mean age of the patients was 55.2 years, with 16.6% aged <50 years. The mean A1C was 7.48% for all patients, 8.32% for those receiving CGRs, and 7.17% for those not receiving CGRs. The NCQA denominator included an additional 74,531 (36.7%) individuals with at least one serious medical or mental health condition, including psychoses or substance abuse (23.5%), conditions with limited life expectancy (6.2%), advanced renal or retinal complications of diabetes (4.5%), serious neurological conditions (4.2%), and serious medical conditions (6.9%).

The A1C distributions were similar for the NCQA and non-CGR denominators but differed markedly for the CGR denominator (Fig. 1). Of the patients not receiving CGRs, 56.0% had A1C values <7%, 12.6% had A1C values of 7.0–7.4%, and 8.8% had AIC values of 7.5–7.9%, compared with 24.8, 12.2, and 12.2% for the patients receiving CGRs. Compared with the non-CGR denominator, in which 22.6% had A1C values >7.9%, there was increased skewness of the NCQA and CGR denominators, for which the proportions were 30.1 and 50.7%, respectively.

At the facility level, the mean and range of percent achievement of a <7% threshold varied considerably between the NCQA and CGR denominators for each ranking category (Table 2). For example, the top 10% performing facilities overall achieved a rate of 57% (range 54–62%) compared with 34% (31–38%) for individuals receiving CGRs. The use of a continuous measure increased the percent achievement compared with the dichotomous measure by 12% for both the overall and CGR denominator. There was an overall 1.2% lower achievement for the denominator that used additional exclusion criteria (data not shown).

Table 1—*Characteristics of NCQA denominator, CGR denominator, and NCQA denominator with additional exclusion criteria*

Attribute (FY 2004)	NCQA denominator	Non-CGR denominator	CGR denominator	NCQA denominator, additional exclusions
Total *N*	203,302	148,951	54,351	128,771
Mean *n* per facility (SD)	1,601 ± 933	1,173 ± 688	428 ± 252	1,014 ± 599
Range	480–5,411	347–4,176	109–1,490	276–3,553
Age (years)	55.2 ± 6.3	55.5 ± 6.1	54.4 ± 6.8	55.4 ± 6.4
<50 years (%)	16.6	15.4	20.1	15.8
50–64 years (%)	83.4	84.6	79.9	84.2
Sex (%)				
Female	4.9	4.9	4.6	4.6
Male	95.2	95.1	95.4	95.4
Veterans' priority (%)				
Missing	1.2	1.3	0.8	1.4
>50% SC disability	25.0	22.7	31.2	19.5
30–40% SC disability	10.0	9.9	10.1	11.1
Other disability, POW, etc.	14.2	14.8	12.6	16.3
Catastrophic disability	2.9	2.4	4.1	1.1
Poverty	30.1	30.7	28.7	29.9
Other	1.5	1.7	0.7	1.8
Copay	15.2	16.4	11.8	18.9
Marital status (%)				
Married	55.5	56.3	53.4	60.4
Other	44.5	43.8	46.6	39.6
Medication profile (%)				
No medications	32.4	44.2	NA	35.6
OHA only	45.7	55.9	17.8	46.2
≥3 OHAs	4.8	NA	17.8	5.2
1 or 2 OHAs	40.9	55.9	NA	41.0
OHA and insulin	13.9	NA	52.0	11.9
Insulin only	8.1	NA	30.2	6.3
A1C level (%)	7.48 ± 1.83	7.17 ± 1.66	8.32 ± 1.99	7.50 ± 1.77
Decreased life expectancy (%)	6.2	6.1	6.7	NA
Advanced complications (%)	5.0	3.3	9.5	NA
Serious medical conditions (%)	6.9	5.5	10.7	NA
Serious neurological conditions (%)	3.7	3.4	4.6	NA
Serious mental health conditions (%)	23.5	22.3	26.9	NA

Data are means ± SD or %. OHA, oral hypoglycemic agent; POW, prisoner of war; SC, service connection; NA, not available.

The overall correlation of facility rankings using the NCQA measure with the CGR denominator was 0.61 and in a sensitivity analysis was 0.65 for A1C <7.5% and 0.74 for A1C <8%. The correlation was 0.97 with the additional exclusion criteria denominator. Rankings of both the aggregate and CGR denominators using the dichotomous threshold were comparable to those using a continuous measure (0.96 and 0.90).

However, the identification of best and worst facilities differed markedly between the NCQA denominator (all patients) and the CGR denominator (Table 3). In this comparison, 8 of 24 facilities that were statistically different identified as being best (five stars) or worst (one star) using the NCQA measure became not significantly different (three stars), and an additional 5 of 24 moved up in ranking. Overall, 42 of 68 (62%) facilities ranked statistically better or worse using the NCQA denominator were ranked as average (not statistically different) using the CGR denominator; 12 of 59 (20%) ranked average became significantly higher or lower. Six of the 31 four- or five-star ranked facilities were ranked lower when exclusion criteria for significant comorbid conditions were applied.

CONCLUSIONS — We demonstrated that overall rankings of VHA facilities based on achievement of the NCQA-HEDIS <7% A1C measure as specified were only modestly correlated with rankings for the subset of patients receiving CGRs. When evaluated based on the NCQA-HEDIS practice of awarding "stars" to denote plan quality rankings, approximately two-thirds of facilities ranked best (four or five stars) overall and about 6 of 10 of those ranked worst (one or two stars) on the current measure were ranked average (three stars, not statistically different) when evaluated using the CGR denominator.

Although it is possible that the observed difference in rankings between the NCQA measure and the CGR denominator may reflect differences in quality of care provided to patients receiving more complex treatment regimens, it could also reflect patient characteristics among facilities. For example, in the current study, the facility percentage of patients receiving CGRs varied twofold (17.9 to 35.2%). In a prior study (16), the 10th and 90th percentile of VA facilities differed mark-
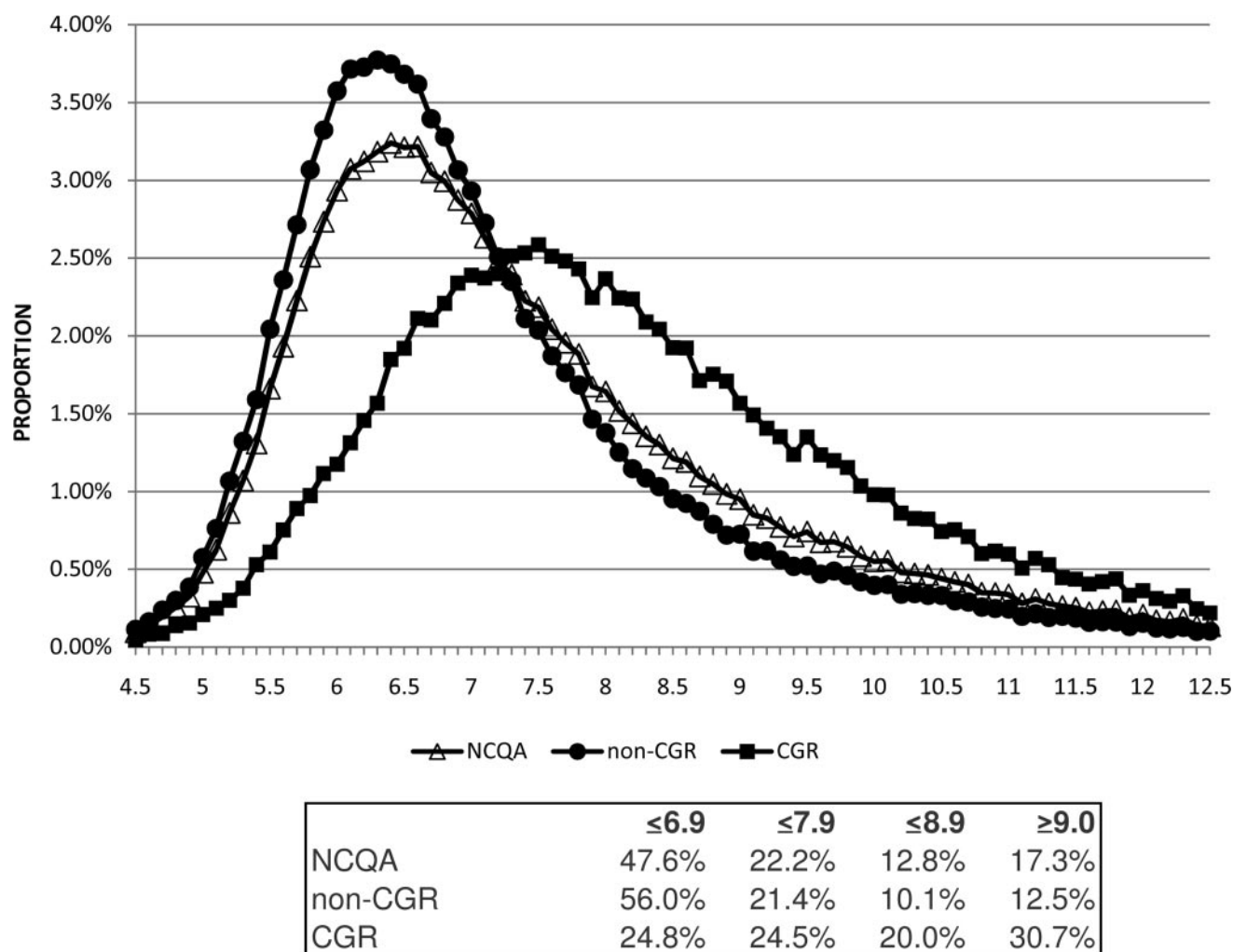
Figure 1—*The percentages of A1C values occurring in 0.1% intervals are plotted separately for each of the denominators. The percentage within each nonoverlapping A1C range within each denominator is reported. Relatively similar distributions are observed for the NCQA (△) and non-CGR(●), whereas the CGR (■) distribution differed more markedly in both skewness and kurtosis as a result of a greater proportion of patients having higher A1C values.*

edly for the proportion of patients with a duration of diabetes >11 years (30.1 and 41.8%), no high school diploma (18.9 and 44.2%), concern over food sufficiency (9.3 and 19.5%), and BMI >35 (11.2 and 19.9%) (16). Facilities with a greater proportion of patients with complex treatment regimens may be at a distinct disadvantage in achieving a <7% threshold.

Our results are relevant to organizations that develop and endorse diabetes measures as well as stakeholders who rely upon them for inferences regarding quality of care. Individuals in whom therapy with two oral agents has failed may be reluctant to start insulin therapy because of valid concerns regarding the benefits, risks, costs, and time requirements of insulin treatment (8). However, thiazolidinediones have demonstrated only

limited efficacy as third-line agents in achieving and maintaining A1C <7% (17), and the long-term efficacy and safety of newer classes of antiglycemic medications as third-line therapy are not yet known.

Furthermore, only about one-third of highly selected patients started on basal insulin (4,18) and about half using a combination of long-acting insulin and prandial insulin (3) can achieve and maintain A1C <7% over 26–52 weeks. Frequent or severe hypoglycemia may be neither acceptable nor safe for many patients, especially those with comorbid conditions and less willingness or ability to comply with recommended self-management. Therefore, the feasibility of achieving <7% A1C in individuals with diabetes of longer duration in general clinical practice is not known.

Consequently, reliance on a <7% threshold measure as the "quality standard" for public reporting or pay for performance could have the unintended consequence of adversely affecting patient safety. Recent trials have demonstrated the potential risks of serious hypoglycemia even in patients who cannot achieve a <7% target (6), and there are significant costs associated with insulin-related hypoglycemia even among employed individuals (19). Moreover, the true value of a single A1C test result of 7% from an accredited laboratory can be within a 6.65 to 7.35% range, assuming a permissible intra-assay coefficient of variation of 5% (20). Point-of-care assays are even more variable, yet A1C testing in the physician office is exempt from external oversight (20). Clinical decisions to initiate or intensify insulin need to be made on

Table 2—*Achievement of A1C threshold and continuous measure within ordinal ranks using NCQA, additional comorbid condition exclusions and CGR denominators*

	NCQA star ranking					
	5 stars	4 stars	3 stars	2 stars	1 stars	Total
NCQA denominator measured using <7% threshold: No. facilities	12	19	59	25	12	127
Proportion <7% minimum (%)	54	51	45	41	34	34
Proportion <7% maximum (%)	62	54	52	46	41	62
Mean proportion <7% (%)	57	53	48	44	39	48
NCQA denominator with additional exclusions removed	12	17	67	19	12	127
Proportion <7% minimum (%)	53	50	43	40	33	33
Proportion <7% maximum (%)	62	53	51	45	40	62
Mean proportion <7% (%)	57	51	47	42	37	47
NCQA denominator measured using continuous measure	12	21	57	25	12	127
Proportion achieved: minimum (%)	66	62	57	54	46	46
Proportion achieved: maximum (%)	75	66	64	58	54	75
Mean proportion (%)	69	64	60	56	52	60
CGT denominator measured using <7% threshold	12	4	89	10	12	127
Proportion <7% minimum (%)	31	29	19	19	13	13
Proportion <7% maximum (%)	38	30	30	21	18	38
Mean proportion <7% (%)	34	30	25	20	17	25
CGR denominator measured using continuous measure	12	8	85	10	12	127
Proportion achieved: minimum (%)	44	41	31	31	28	28
Proportion achieved: maximum (%)	51	44	41	33	30	51
Mean proportion (%)	47	42	37	32	30	37

Data are *n* or %. Within each denominator, facilities were ranked within five ordinal categories based on both *z* score distribution ($\leq$10% [5 stars or best]; 10–33% [4 stars]; 34–66% [3 stars]; 67–90% [2 Stars], and $\geq$90% [1 star or worst]) and statistical significance ($P < 0.05$). The minimum, maximum, and mean proportion of patients meeting the measure are presented for the facilities included within each category.

the basis of careful evaluation of individual patient trends in A1C and self-monitoring results rather than a single A1C value that is marginally higher (i.e., within laboratory error) than the performance measure (21). We propose that rather than assessing "good glycemic control" by an all-or-none threshold, developers should provide credit for an A1C result within an acceptable range to balance the tradeoffs of benefits, harms, and patient preferences.

Alternatively, a continuous measure that can assess incremental improvement toward an A1C <7% target for patients receiving CGRs may have significant desirable attributes over an all-or-none threshold measure with respect to relevance, scientific soundness, and feasibility. At the population level, it would provide more information for inferential hypothesis testing compared with a categorical or binomial measure (22) and could minimize statistical variation at a population level due to either A1C measurement accuracy and bias (20) or seasonal variation (23). It is also consistent with the evolving quality improvement literature that demonstrates the difficulty in increasing attainment of <7% A1C, especially in patients with starting A1C levels >8%, despite structured interventions in staff model practice settings with electronic health care records (24). Because it is not possible to identify type 1 diabetes or duration of disease using administrative data, stratification by treatment regimen is a pragmatic approach that can minimize attribution error; that is, better glycemic control is incorrectly interpreted as improved plan quality when it actually reflects a greater proportion of patients with recent-onset diabetes (patients not receiving CGRs) in the denominator.

In addition, from both an individual and a population health perspective, there is a greater health benefit from treating higher A1C values (>8%) than values marginally >7%. A continuous measure can better reward the additional time and costs of a multidisciplinary team approach spent on intensive therapy of patients with complex treatment regimens even if they do not reach the <7% "goal." This strategy could incentivize physician practices to compete for referrals or enter into contractual relationships with endocrinologists or disease management programs, thus focusing scarce resources on the most difficult patients with the greatest need.

Finally, although exclusion criteria had a minimal impact on rankings, we suggest that performance measure developers incorporate additional exclusion criteria that would be consistent with guideline recommendations (7). Because the current measure excludes individuals with a single cardiovascular ICD9-CM code regardless of therapy, it seems reasonable to either exclude or stratify patients with comorbid conditions that might increase the risk of serious hypoglycemia with CGRs. Further study is needed to define the sensitivity and specificity of algorithms to identify such patients.

Our study is limited by the ability to assess severity of chronic disease or patient level factors using administrative data. We also did not determine the starting point for glycemic control, medication persistence, or adverse events (including hypoglycemia). However, these are general limitations of using only administrative data for evaluation of industry standard cross-sectional measures. Additional studies are necessary to ascertain whether our findings in the VHA are generalizable to other health care systems.

In summary, the assessment of the quality of good glycemic control among

Table 3—*Changes in facility rankings compared with NCQA standard measure*

	NCQA rating					
	5 stars	4 stars	3 stars	2 stars	1 stars	Total
NCQA denominator <7% measure	12	19	59	25	12	127
NCQA denominator with additional comorbid conditions excluded						
5 stars	9	3				12
4 stars	3	13	1			17
3 stars		3	58	5	1	67
2 stars				19		19
1 star				1	11	12
NCQA denominator using a continuous measure						127
5 stars	10	2				12
4 stars	1	14	6			21
3 stars	1	3	50	3		57
2 stars			3	20	2	25
1 star				2	10	12
CGR denominator using <7% threshold						
5 stars	6	2	4			12
4 stars		3	1			4
3 stars	6	14	47	20	2	89
2 stars			1	4	5	10
1 star			6	1	5	12
CGR denominator using continuous measure						
5 stars	6	3	3			12
4 stars	1	3	4			8
3 stars	5	13	48	18	1	85
2 stars				6	4	10
1 star						

Changes in facility rankings, compared with the NCQA-specified measure, are presented for a denominator that excludes patients with additional serious comorbid conditions and a denominator of patients receiving CGRs, using both threshold and continuous measures.

VHA facilities differs using the NCQA-HEDIS measure compared with a subset of patients receiving complex glycemic treatment regimens. Our recommendation to stratify measures by CGR and to award partial credit using a continuous measures are consistent with proposals made at a 2006 AHRQ funded conference (25). Although the use of a single number for all enrollees to reflect the quality of glycemic treatment within a plan may be attractive in its simplicity, it is superficial and may be misleading. A lack of an appreciation of the hidden complexities of care and measurement by policy makers and managers may thus result in actions that have unintended and potentially harmful consequences for patients, providers, plans, and society.

References
1. National Committee for Quality Assurance. *HEDIS 2009 Technical Specifications*. Vol. 2. Washington, DC, National Committee for Quality Assurance, 2009
2. Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). UK Prospective Diabetes Study (UKPDS) Group. Lancet 1998;352:837–853
3. Holman RR, Thorne KI, Farmer AJ, Davies MJ, Keenan JF, Paul S, Levy JC, 4-T Study Group. Addition of biphasic, prandial, or basal insulin to oral therapy in type 2 diabetes. N Engl J Med 2007;357:1716–1730
4. Horvath K, Jeitler K, Berghold A, Ebrahim SH, Gratzer TW, Kaiser T, Plank J, Pieber TR, Siebenhofer A. Long-acting insulin analogues versus NPH insulin (human isophane insulin) for type 2 diabetes mellitus. Cochrane Database Syst Rev 2007;18:CD005613
5. Bainbridge KE, Cowie CC, Rust KF, Fradkin JE. Mitigating case mix factors by choice of glycemic control performance measure threshold. Diabetes Care 2008;31:1754–1760
6. Bonds DE, Miller ME, Bergenstal RM, Buse JB, Byington RP, Cutler JA, Dudl RJ, Ismail-Beigi F, Kimel AR, Hoogwerf B, Horowitz KR, Savage PJ, Seaquist ER, Simmons DL, Sivitz WI, Speril-Hillen JM, Sweeney ME. The association between symptomatic, severe hypoglycaemia and mortality in type 2 diabetes: retrospective epidemiological analysis of the ACCORD study. BMJ. 8 January 2010 [Epub ahead of print]
7. Qaseem A, Vijan S, Snow V, Cross JT, Weiss KB, Owens DK, Clinical Efficacy Assessment Subcommittee of the American College of Physicians. Glycemic control and type 2 diabetes mellitus: the optimal hemoglobin A1c targets. A guidance statement from the American College of Physicians. Ann Intern Med 2007;147:417–422
8. Peyrot M, Rubin RR, Lauritzen T, Skovlund SE, Snoek FJ, Matthews DR, Landgraf R, Kleinebreil L. Resistance to insulin therapy among patients and providers: Results of the cross-national Diabetes Attitudes, Wishes, and Needs (DAWN) study. Diabetes Care 2005;28:2673–2679
9. Kupersmith J, Francis J, Kerr E, Krein S, Pogach L, Kolodner RM, Perlin JB. Advancing evidence-based care for diabetes: lessons from the Veterans Health Administration. Health Affairs 2007;26:w156–w168
10. Pogach LM, Rajan M, Aron DC. Aligning performance measurement with clinical epidemiology: comparison of weighted

performance measurement and dichotomous threshold for glycemic control in the Veterans Health Administration. Diabetes Care 2006;29:241–246

11. Miller DR, Safford MM, Pogach LM. Who has diabetes? Best estimates of diabetes prevalence in the Veterans Health Administration based on computerized patient data. Diabetes Care 2004;27(Suppl. 2): B10–B21

12. Sales MM, Cunningham FE, Glassman PA, Valentino MA, Good CB. Pharmacy benefits management in the Veterans Health Administration: 1995 to 2003. Am J Manag Care 2005;11:104–112

13. Pogach LM, Tiwari A, Maney M, Rajan M, Miller DR, Aron DC. Should mitigating co-morbidities be considered in assessing healthcare plan performance In achieving optimal glycemic control. Am J Manag Care 2007;13:133–140

14. Hofer TP, Hayward RA, Greenfield S, Wagner EH, Kaplan SH, Manning WG. The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. JAMA 1999;281:2098–2105

15. State of Texas Office of Public Insurance Counsel and the Department of State Health Services Center for Health Statistics. *Guide to Texas HMO Quality: 2008.* Available from http://www.opic.state. tx.us/docs/544_guidetohmoquality08.pdf. Accessed 12 March 2010

16. Maney M, Tseng CL, Safford MM, Miller DR, Pogach LM. Impact of self-reported patient characteristics upon assessment of glycemic control in the Veterans Health Administration. Diabetes Care 2007;30: 245–251

17. Tran MT, Navar MD, Davidson MB. Comparison of the glycemic effects of rosiglitazone and pioglitazone in triple oral therapy in type 2 diabetes. Diabetes Care 2006;29:1395–1396

18. Rosenstock J, Dailey G, Massi-Benedetti M, Fritsche A, Lin Z, Salzman A. Reduced hypoglycemia risk with insulin glargine: a meta-analysis comparing insulin glargine with human NPH insulin in type 2 diabetes. Diabetes Care 2005;28:950–955

19. Rhoads GG, Orsini LS, Crown W, Wang S, Getahun D, Zhang Q. Contribution of hypoglycemia to medical care expenditures and short-term disability in employees with diabetes. J Occup Environ Med 2005;47:447–452

20. Holmes EW, Erçsahin C, Augustine GJ, Charnogursky GA, Gryzbac M, Murrell JV, McKenna KM, Nabhan F, Kahn SE. Analytic bias among certified methods for the measurement of hemoglobin A1c: a cause for concern? Am J Clin Pathol 2008; 129:540–547

21. Aron DC, Pogach LM. One size does not fit all: the need for a continuous measure for glycemic control in diabetes. Jt Comm J Qual Patient Saf 2007;33:636–643

22. Burnham KP, Anderson DR. *Model Selection and Multi-Model Inference: A Practical Information-Theoretic Approach.* 2nd ed. New York, Springer, 2002

23. Tseng CL, Brimacombe M, Xie M, Rajan M, Wang H, Kolassa J, Crystal S, Chen TC, Pogach L, Safford M. Seasonal patterns in monthly hemoglobin A1c values. Am J Epidemiol 2005;161:565–574

24. Shojania KG, Ranji SR, McDonald KM, Grimshaw JM, Sundaram V, Rushakoff RJ, Owens DK. Effects of quality improvement strategies for type 2 diabetes on glycemic control: a meta-regression analysis. JAMA 2006;296:427–440

25. Kerr EA. Assessing Quality of Care for Diabetes [article online], 2008. Available from http://www.ahrq.gov/qual/ diabetescare/diabetescare.pdf. Accessed 12 March, 2010