Performance Status of Health Care Facilities Changes With Risk Adjustment of HbA_{1c}

QUANWU ZHANG, PHD
MONIKA SAFFORD, MD
JOHN OTTENWELLER, PHD
GERALD HAWLEY, MSN
DENIS REPKE, PHD
JAMES F. BURGESS JR., PHD
SUNIL DHAR, PHD
HSIAOFEN CHENG, MS
HERBERT NAITO, PHD
LEONARD M. POGACH, MD, MBA

OBJECTIVE — To develop a risk adjustment method for HbA_{1c} based solely on administrative data and to determine the extent to which risk-adjusted HbA_{1c} changes the identification of high- or low-performing medical facilities.

RESEARCH DESIGN AND METHODS — Through use of pharmacy records, 204,472 diabetic patients were identified for federal fiscal year 1996 (FY96). Complete information (HbA_{1c} levels, demographic data, inpatient records, outpatient pharmacy utilization records) was available on 38,173 predominantly male patients from 48 Veterans Health Administration (VHA) medical facilities. Hierarchical mixed-effects models were used to estimate risk-adjusted unique facility–level HbA_{1c}.

RESULTS — Predicted HbA_{1c} demonstrated expected patterns for major factors known to influence glycemic control. Poorer glycemic control was seen in minorities and patients with greater disease severity, longer duration of disease (using treatment type or presence of amputation as surrogates), and more extensive comorbidity (measured by an adapted Charlson index). Better glycemic control was seen in Caucasians, older diabetic patients, and patients with higher outpatient utilization. The number of performance outliers was reduced as a result of risk adjustment. For mean HbA_{1c} levels, 7 facilities that were initially identified as statistically significant outliers were no longer outliers after risk adjustment. For high-risk HbA_{1c} (>9.5%) rates, 12 facilities that were initially identified as statistically significant outliers were no longer outliers after risk adjustment.

CONCLUSIONS — Risk adjustment using only administrative data resulted in substantial changes in identification of high or low performers compared with non–risk-adjusted HbA_{1c}. Although our findings are exploratory, risk adjustment using administrative data may be a necessary and achievable step in quality assessment of diabetes care measured by rates of high-risk HbA_{1c} (>9.5%).

Diabetes Care 23:919–927, 2000

From the Veterans Affairs Medical Center (Q.Z., M.S., J.O., L.M.P.), East Orange; the Departments of Neurosciences (Q.Z., J.O.) and Medicine (M.S., L.M.P.), University of Medicine and Dentistry of New Jersey, New Jersey Medical School, Newark; the Department of Applied Mathematic Sciences (S.D., H.C.), New Jersey Institute of Technology, Newark, New Jersey; the Veterans Affairs Health Care Analysis Information Group (G.H., D.R.), Milwaukee, Wisconsin; the Veterans Affairs Management Science Group (J.F.B.), Bedford, Massachusetts; the Boston University School of Public Health (J.F.B.), Boston, Massachusetts; and the Department of Clinical Chemistry (H.N.), Cleveland Veterans Affairs Medical Center, Cleveland, Ohio. (Q.Z. is currently affiliated with ITG, Schering-Plough Corporation, Kenilworth, New Jersey; H.C. is currently affiliated with DoubleClick, Inc., New York, New York)

Address correspondence and reprint requests to Leonard M. Pogach, MD, MBA, Endocrinology, Veterans Affairs Medical Center, 385 Tremont Ave., East Orange, NJ 07018. E-mail: leonard.pogach@med.va.gov.

Received for publication 30 June 1999 and accepted in revised form 1 March 2000.

Abbreviations: DQIP, Diabetes Quality Improvement Project; FY96, fiscal year 1996; HEDIS, Health Plan Employer Data and Information Set; ICD-9-CM, *International Classification of Diseases, 9th Revision, Clinical Modification*; NCCC, National Center for Cost Containment; NHANES 3, Third National Health and Nutrition Examination Survey; VHA, Veterans Health Administration.

A table elsewhere in this issue shows conventional and Système International (SI) units and conversion factors for many substances.

High HbA_{1c} is a proven risk factor for the development of microvascular complications and poor quality of life in individuals with diabetes (1–4). HbA_{1c} levels are therefore a key indicator for assessing the quality of diabetes care in populations. The Diabetes Quality Improvement Project (DQIP), a federal private-sector coalition, has recommended that HbA_{1c} be included as an outcome measure to promote accountability, performance comparison, and quality improvement (5).

However, HbA_{1c} is subject to well-known influences outside of the control of health care providers or health plans; such influences (which include include age, race/ethnicity, type and duration of diabetes, comorbidity, and patient adherence) should be accounted for to avoid unfair comparisons. Unfortunately, many of these risk factors are not readily available from administrative data sources. Developing effective risk adjustment for performance measurement poses particular challenges, and no established risk adjustment model for HbA_{1c} currently exists.

Consequently, DQIP conservatively established the threshold for "poor glycemic control" at HbA_{1c} >9.5%, assuming that adjusting for risks and lack of laboratory standardization would be less important with this relatively high measure. It is uncertain to what extent risk adjustment of this chosen threshold may be warranted even after nationwide implementation of the National Glycohemoglobin Standardization Program (6–9).

We developed a risk adjustment model using only administrative data in the Veterans Health Administration (VHA), the nation's largest integrated health care system. Using facility-specific automated data systems and the central VHA database repository, a large diabetes registry linking inpatient discharges and demographic, pharmacy, and laboratory files was made available for our study (10). Specifically, we examined the extent to which performance status of medical facilities (characterized by statistically significant differences from the average VHA HbA_{1c}) changed with risk adjustment.

RESEARCH DESIGN AND METHODS

Diabetes case identification

Using a pharmacy database, veterans with diabetes were identified as those who had received insulin, oral agents (metformin, sulfonylurea, or acarbose), or blood glucose monitoring supplies from a VHA medical facility during federal fiscal year 1996 (FY96) (from 1 October 1995 to 30 September 1996) (10).

Setting and data

During the study period, the VHA comprised 173 medical facilities and 401 ambulatory clinics, providing care to 2.6 million veterans. Information on diabetic patients was collected through a program initiated by the VHA National Center for Cost Containment (NCCC) to determine the prevalence, costs, and outcomes of diabetes. Reporting in 1996 was voluntary; 65 facilities were thus included.

From pharmacy files, we identified 204,472 diabetic patients with 1,017,873 records from the 65 facilities (10); 80,809 patients had at least 1 laboratory value. There were 103,459 inpatient records on 59,598 individuals with up to 10 *International Classification of Diseases, 9th Revision, Clinical Modification* (ICD-9-CM) codes. The total number of inpatient records included 2,219 records on 1,925 individuals hospitalized with incident amputation, and 8,670 records on 6,187 individuals hospitalized for lower-extremity ulcers. Also available were residential postal codes (zip codes) for each of the 204,472 identified patients and normative ranges of HbA_{1c} for each reporting facility.

Of the 80,809 patients with at least 1 laboratory value, 53,361 (66%) were found to have had at least 1 true HbA_{1c} result (those with glycohemoglobin were excluded). After excluding patients with missing or invalid values on any of the predictor variables, we included 38,173 of the 53,361 (71.5%) patients from 48 facilities in our analyses. We compared this reduced sample with the original cohort and found no difference in the distributions of age and sex. We did find, however, that the reduced sample significantly underrepresented white Hispanic and African-American patients and overrepresented Caucasian patients. Nevertheless, neither the mean HbA_{1c} level (2.199) nor the high-risk rate (0.28) in the reduced sample differed substantially from those in the whole sample (2.181 and 0.26, respectively).

Measures

The upper limit of normative ranges of HbA_{1c} varied between 5 and 8.2% across facilities. To reduce the impact of differences in HbA_{1c} measurement methodology (11,12), we calculated upper-limit–adjusted HbA_{1c} (HbA_{1c} level) by subtracting the upper limit of the normative range at each facility from individual HbA_{1c} values. The dichotomous measure of HbA_{1c} used 9.5% as a cutoff, defined as high-risk HbA_{1c} or high-risk rate (5).

In constructing our risk adjustment model, we sought to control for those extrinsic factors known to influence HbA_{1c} that are likely to be outside the control of a health system (13). Because of the nature of administrative data, we faced several challenges. First, aside from demographic variables, most measures related to diabetes care have both extrinsic and intrinsic components or implications. For example, treatment modality may be considered an intrinsic variable because it is determined, at least in part, by physicians. Treatment choice, however, is complex and determined by multiple factors, including diabetes type, duration (an extrinsic factor with which treatment choice has been strongly correlated [14]), medical history, the patient's treatment preferences, as well as the physician's practice patterns.

Second, a central issue for this study was to use as much information as available in the data to achieve operationalization of such key concepts as diabetes severity, diabetes duration, and health services utilization. Without survey data or complete physiological measures, some of these concepts had to be approximated by using surrogate administrative variables. We constructed 4 general types of measurement factors: demographic; health care seeking; geographic; and diabetes severity, duration, and comorbidity.

Demographic factors included age, sex, and race/ethnicity. Age and race/ethnicity are well-documented correlates of HbA_{1c} (15). To approximate access to care, we included an eligibility (priority) rating for VHA medical care. Eligibility classification for a patient was administratively determined; a higher percentage of connection to military service was associated with higher eligibility to receive VHA medical care during the study period.

Health care–seeking factors were also included. Diabetes is an illness that requires patient compliance for successful treatment; it requires patients to adhere to prescribed treatments and to follow appropriate referral recommendations to subspecialty or ancillary services. We thus included total pharmacy costs (patient expenditures on diabetes-specific drugs and supplies) and number of outpatient contacts to a VHA facility, including all facility contacts, not only office visits (as reported by VHA local facilities). These 2 variables were log-transformed because of substantial variability.

Geographic factors are reported to be predictive of utilization of recommended procedures for diabetic patients (16). Because the quality of care for common medical conditions has been reported to be lower in rural hospitals (17,18), we identified metropolitan areas in which patients lived during FY96 by using their residential zip codes as an additional measure.

Duration and severity of diabetes are central measures to quantify risks for outcomes of diabetes care. Diabetes treatment modality has been found to be highly correlated with duration of type 2 diabetes. The Third National Health and Nutrition Examination Survey (NHANES 3) showed that insulin use increased rapidly and oral agent use decreased within 5 years of diagnosis (14). Likewise, in the U.K. Prospective Diabetes Study, most patients progressed to insulin therapy despite the closely monitored setting of a clinical trial (2). Because >90% of people with diabetes in the general population have type 2 diabetes (10,19), we used treatment modality as a proxy measure for diabetes duration. To control for severity of disease, we included hospitalization for lower-extremity amputation or ulcers. Home glucose monitoring was included as a potential correlate of both disease severity and compliance (14).

The Charlson index, originally developed to classify prognostic comorbidity in longitudinal studies, has been adapted for use in predicting general mortality risk using administrative data (20,21). Hospital discharges, including up to 10 discharge ICD-9-CM diagnoses, were classified into 0–6 categories according to severity and the extent of comorbid conditions. Diabetes-related diagnoses were by definition not part of comorbidity and were thus excluded. Multiple diagnoses were summed up; i.e., a diagnosis was only counted once if duplicate diagnoses existed within FY96. Because outpatient diagnostic information was not available in FY96, all patients without hospitalization received a 0 score for comorbidity. This measure

Table 1—*Descriptive statistics, numbers of Veterans Affairs outpatient contacts, and simple group comparisons of HbA_{1c} levels and rates (%) of high-risk HbA_{1c}*

	n	No. of contacts	HbA_{1c} level* (95% CI)	High-risk HbA_{1c}	
				Rates (95% CI)	Odds ratios (*P*)
Age (years)					
<45	1,897	15.36	3.0 (2.90–3.12)	0.43 (0.41, 0.46)	1.67 (<0.01)
45–54	6,039	18.05	2.6 (2.57–2.68)	0.37 (0.36, 0.37)	1.29 (<0.01)
55–64	8,848	16.39	2.3 (2.33–2.42)	0.32 (0.31, 0.32)	—†
65–74	15,112	15.83	2.0 (2.00–2.05)	0.24 (0.24, 0.25)	0.69 (<0.01)
>74	6,278	15.80	1.7 (1.68–1.76)	0.19 (0.18, 0.20)	0.51 (<0.01)
ANOVA		$F_{(4, 38169)} = 18.15$ $P < 0.0001$	$F_{(4, 38169)} = 274.61$ $P < 0.0001$		
Sex					
Female	723	20.20	2.22 (2.06–2.38)	0.27 (0.24, 0.30)	—†
Male	37,451	16.21	2.20 (2.18–2.22)	0.28 (0.28, 0.28)	1.05 (0.59)
ANOVA		$F_{(1, 38172)} = 32.48$ $P < 0.0001$	$F_{(1, 38172)} = 0.07$ $P = 0.79$		
Ethnicity/race					
White Hispanic	1,043	21.80	2.46 (2.33–2.59)	0.32 (0.29, 0.35)	1.32 (<0.01)
Black Hispanic	103	20.88	2.31 (1.95–2.68)	0.33 (0.30, 0.36)	1.37 (0.13)
Native American	127	18.14	2.27 (1.83–2.71)	0.31 (0.27, 0.36)	1.28 (0.19)
African-American	6,860	18.17	2.47 (2.41–2.91)	0.35 (0.34, 0.36)	1.49 (<0.01)
Asian	70	26.06	2.92 (2.43–3.49)	0.36 (0.24, 0.47)	1.55 (0.08)
Caucasian	29,971	15.62	2.12 (2.10–2.15)	0.26 (0.26, 0.27)	—†
ANOVA		$F_{(5, 38168)} = 45.50$ $P < 0.0001$	$F_{(5, 38168)} = 37.30$ $P < 0.0001$		
Service connection rating (%)					
50–100	6,241	22.21	2.14 (2.09–2.19)	0.26 (0.25, 0.27)	0.91 (<0.01)
<50	5,771	15.21	2.31 (2.25–2.36)	0.30 (0.29, 0.31)	1.08 (0.02)
0	26,162	15.11	2.19 (2.16–2.21)	0.28 (0.28, 0.29)	—†
ANOVA		$F_{(2, 38171)} = 383.5$ $P < 0.0001$	$F_{(2, 38171)} = 10.96$ $P < 0.0001$		

*Adjusted for the upper limits of normative ranges (see text); †reference category used in logistic regression for simple comparisons of the rates of high-risk HbA_{1c}.

therefore only represented additional risk adjustment among inpatients.

Analysis

We used mixed-effects models to estimate HbA_{1c} for several reasons. First, we were interested in glycemic control at the facility level, but data were collected from individuals. Traditional methods of aggregation by pooling individual data into facility-level measurements have generated concerns about aggregation bias (16,22). Additionally, patients were nested within participating medical facilities unevenly (23). Traditional analytic approaches would yield less reliable estimates of HbA_{1c} from facilities with fewer patients compared with those with more patients.

We proceeded with the analyses by taking 4 steps to disentangle and identify the potential roles of each cluster of risk factors in predicting both HbA_{1c} level and high-risk rate. First, we constructed model 1 using age, sex, race/ethnicity, and degree of military service connection. This model established a baseline prediction, adjusting for patient demographic characteristics that were entirely exogenous to HbA_{1c}; it would not be affected by any disease or treatment processes. Next, in model 2, we added pharmacy spending, outpatient contacts, and geographic location to examine whether health care seeking or utilization made additional contributions. Next, in model 3, we added proxy measures for diabetes severity and duration: hospitalization for lower-extremity amputation or ulceration, treatment modality, blood glucose monitoring, and length of stay. Without controlling for these variables, the effects of more frequent contacts and higher pharmacy spending in model 2 were expected to partially reflect severity of diabetes. In model 4, we included the Charlson comorbidity score to determine the extent to which the previous models were affected by comorbid conditions, as a way to assess the sensitivity of our risk adjustment results.

To evaluate the proportion of variation in HbA_{1c} potentially attributable to institutional factors apart from individual factors, we estimated intraclass correlations (22–24). An intraclass correlation is the proportion of facility-level variability relative to the total variability in HbA_{1c} among individuals and among facilities; that is, the maximum amount of variability that could have been explained by facility-level predictors. The intraclass correlations indicated (at most) a 10–11% variability in HbA_{1c} levels and high-risk rates due to facility-level factors.

Individual HbA_{1c} levels were adjusted by risk factors; unique facility-level HbA_{1c} values were estimated simultaneously. Deviations of facility-specific HbA_{1c} estimates from the estimated VHA national average were tested for both the high-risk HbA_{1c} rates and mean HbA_{1c} levels. Facilities that were statistically ($P < 0.05$) below or above

Table 2—*Adjusting HbA_{1c} values using hierarchical mixed-effects models*

	HbA_{1c} levels		Rate of high-risk HbA_{1c}	
	Estimate	*P*	Estimate	*P*
Conditional Veterans Affairs average level and rate	1.49	<0.01	0.05	<0.01
Demographic factors				
Age (years)				
<45	0.64	<0.01	0.60	<0.01
45–54	0.33	<0.01	0.33	<0.01
55–64	—	—	—	—
65–74	−0.30	<0.01	−0.33	<0.01
>74	−0.47	<0.01	−0.52	<0.01
Sex				
Male	0.007	0.96	0.13	0.41
Ethnicity/race				
White Hispanic	0.001	0.99	0.14	0.32
Black Hispanic	0.06	0.90	0.03	0.95
Native American	0.17	0.52	0.32	0.29
African American	0.31	<0.01	0.28	<0.01
Asian	1.34	<0.01	0.87	0.06
Caucasian	—	—	—	—
Service connection rating (%)				
50–100	−0.03	0.54	−0.04	0.53
<50	−0.04	0.52	−0.09	0.23
0	—	—	—	—
Care-seeking (utilization) factors				
Pharmacy cost (log)	−0.14	<0.01	−0.14	<0.01
Number of outpatient visits (log)	−0.14	<0.01	−0.11	<0.01
Geographic factor				
Metropolitan areas	−0.04	0.47	−0.05	0.40
Diabetes duration/severity and comorbidity factors				
Amputation or ulceration	0.22	<0.01	0.22	<0.01
Days of hospitalization (log)	−0.08	<0.01	−0.07	<0.01
Treatment modality				
Oral medication	2.17	<0.01	3.16	<0.01
Insulin	1.56	<0.01	2.62	<0.01
Oral and insulin therapies	2.73	<0.01	3.70	<0.01
Diet control only	—	—	—	—
Monitored blood glucose	0.34	<0.01	0.27	<0.01
Charlson index	0.09	<0.01	0.06	0.03

As an illustration for the results presented in this table, for example, a facility with more Caucasian patients who were >64 years of age, controlled by diet alone and with a lower Charlson score, is expected to have a lower mean HbA_{1c} level and lower rate of high-risk HbA_{1c} than a facility with more African-American and younger patients who received insulin therapy.

that average were identified as high- and low-performing outliers, respectively.

RESULTS — The mean age in our final sample was 66 years of age. Within our final sample of 38,173 patients, 31% used insulin, 58% used oral agents, 8% used a combination of insulin and oral agents, and the remaining 3% received only home monitoring supplies and were presumed to be diet-controlled.

Table 1 shows that HbA_{1c} decreased with age monotonically. Minorities had higher levels of HbA_{1c} than Caucasians. Individuals between 45 and 54 years of age had the most frequent contacts at a facility. A military service connection ≥50% was associated with an almost 50% increase in the number of facility contacts and a lower HbA_{1c}. The univariate distributions of patients with a high-risk HbA_{1c} had a pattern similar to the level of HbA_{1c}.

Table 2 shows the results from all of the final models with HbA_{1c} levels and high-risk rates as dependent measures adjusting for all available risk factors. Many estimates confirmed previous reports. For example, age trends were approximately linear for both outcome measures. Relative to Caucasian patients, African-American and Asian patients had higher HbA_{1c}. Hospitalization for lower-extremity amputation or ulcer was associated with poorer glycemic control. Patients treated with oral agents, insulin, or a combination of the 2 had poorer glycemic control than diet-controlled individuals. Those with higher HbA_{1c} were more likely to self-monitor blood glucose.

The 4 steps of risk adjustment allowed us to view the course of some important changes—in effect magnitude and direction—as each set of new variables was added to an existing model. Pharmacy costs initially were associated with poor glycemic control, but after accounting for disease severity and duration resulted in an association with better control. Also, the magnitude of the effects of pharmacy costs and outpatient facility contacts more than doubled when inpatient comorbidity was taken into account, suggesting that these are highly important and correlated variables. A significant negative effect of metropolitan areas for the high-risk HbA_{1c} rates turned nonsignificant when inpatient comorbidity was entered. When comorbidity was controlled for, greater length of hospital stay became a significant predictor of better results of glycemic control.

Although the convenient goodness-of-fit measure of R^2 has no equivalent in a mixed-effects model, it is roughly its lower bound of goodness-of-fit estimates. Other statistics, such as Akaike's information criterion (25), are unbounded and useful only when comparing 2 nested models. Hence, using R^2 as a conservative estimate, our model without random effects explained 10.5% variance of HbA_{1c} levels. Similarly, a logistic model including all risk variables but no random effects yielded a *C* statistic of 0.67 for the high-risk rates.

Figure 1 shows changes in performance status in rates of high-risk HbA_{1c} for the 48 facilities before and after risk adjustment. Five initially low-performing (poorly performing) facilities moved into the average-performing (that is, no longer significantly different) realm with risk adjustment; 6 high-performing facilities moved into the average-performing realm; and 1 average-performing facility moved into the high-

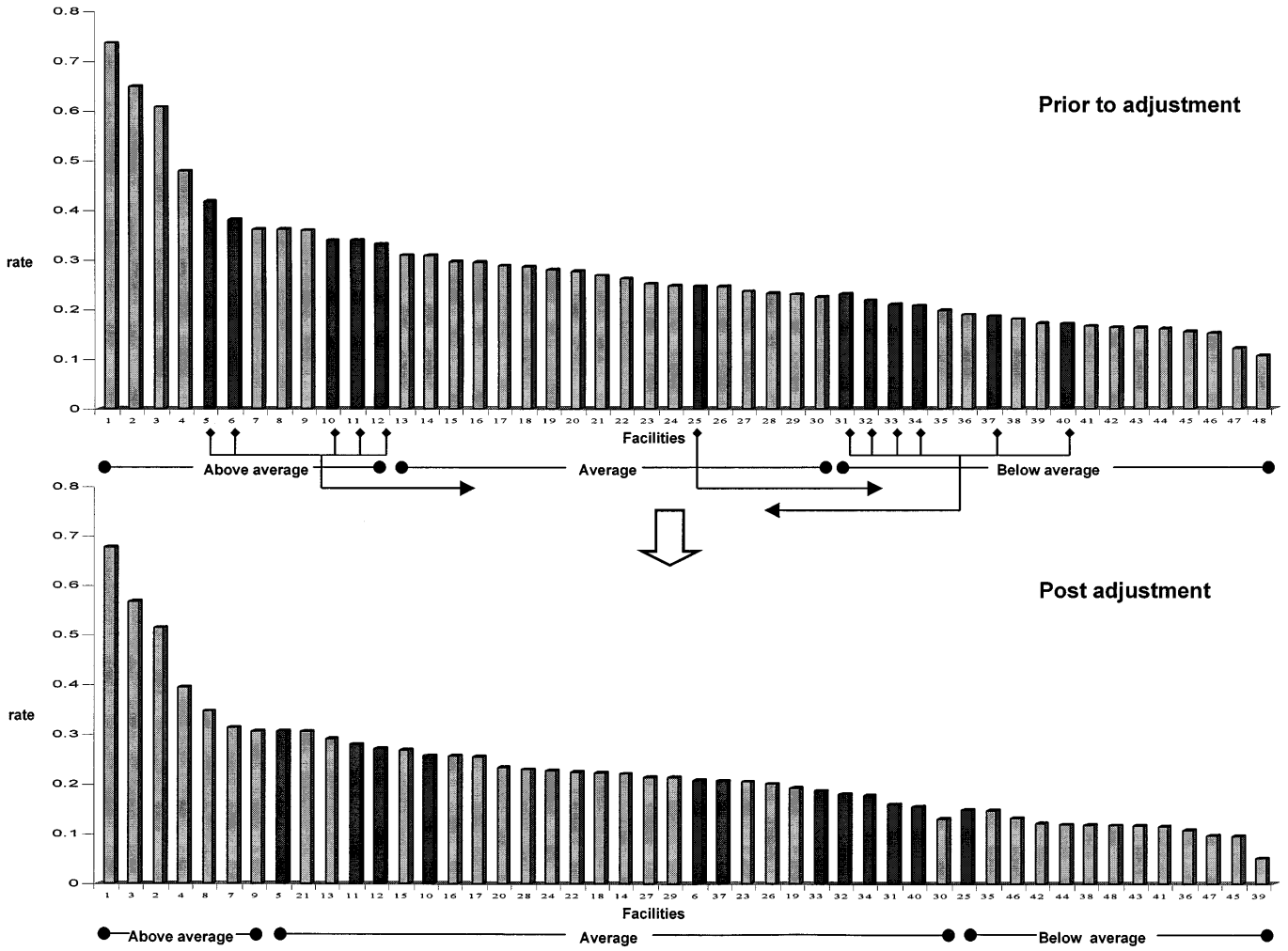

Figure 1—*Changes of performance status with risk adjustment of high-risk rate of HbA_{1c}. "Above average" or "below average" denotes facilities that had a rate of high-risk HbA_{1c} statistically higher or lower than the Veterans Affairs national estimate. Arrows indicate the directions of change in performance status with risk adjustment. Darkened bars represent those facilities whose performance status significantly changed with risk adjustment (e.g., Facility 6 was "above average" before adjustment but relocated near the center of the Veterans Affairs national average after risk adjustment; Facility 37 changed from "below average" to "average" with risk adjustment). Fewer facilities had a rate significantly above or below the Veterans Affairs national mean after risk adjustment as indicated by the expanded range of average performers.*

performing realm. Altogether, 12 of 48 facilities (25%) changed performance status when evaluated based on rates of high-risk HbA_{1c}. Of the 48 facilities, 30 were initially assessed as high-risk HbA_{1c} outliers, which changed to only 20 after risk adjustment. Of note, when evaluated based on mean HbA_{1c} levels, only 7 facilities (15%) changed status (data not shown).

CONCLUSIONS — Diabetes is a major health problem in the U.S., with ample documentation of suboptimal care for many patients, despite strong evidence demonstrating the benefits of good care and the staggering costs this disease incurs (2,16,26–35). DQIP represents a major step toward improvement, with the promise of standardized measures that will permit benchmarking of care across health systems (5). A critical challenge remains, however, in how best to adjust these measures for risks to allow fair comparisons.

Early efforts to identify low-performing hospitals and surgeons of coronary bypass surgery were misleading as a result of lack of risk adjustment (36). Since then, considerable progress in risk adjusting inpatient outcomes has been achieved (37,38), as evidenced, for example, by the state of New York's focus on mortality rates of patients undergoing coronary bypass surgery (39) and Pennsylvania's focus on mortality rates after acute myocardial infarction (40). However, inpatient data sources typically contain elements not readily available in most ambulatory settings. Some valuable strategies from the inpatient experience therefore are not pertinent in the largely outpatient diabetes quality of care setting.

Consequently, DQIP proposed the threshold for high-risk HbA_{1c} to be >9.5% in an attempt to circumvent the lack of an established risk adjustment method for this important diabetes outcome. Our analyses demonstrate that risk adjustment nevertheless changed performance status for a substantial number of facilities. Both the quality improvement (level of HbA_{1c}) and the accountability measures (rate of high-

Appendix 1—*Predicting levels of upper-normal-limit–adjusted HbA_{1c} using a hierarchical mixed-effects model*

	Model 1		Model 2		Model 3		Model 4	
	Estimate	*P*	Estimate	*P*	Estimate	*P*	Estimate	*P*
Conditional Veterans Affairs average level	2.22	<0.01	1.94	<0.01	0.97	<0.01	1.49	<0.01
Demographic factors								
Age (years)								
<45	0.60	<0.01	0.62	<0.01	0.49	<0.01	0.64	<0.01
45–54	0.26	<0.01	0.27	<0.01	0.24	<0.01	0.33	<0.01
55–64	—	—	—	—	—	—	—	—
65–74	–0.34	<0.01	–0.34	<0.01	–0.31	<0.01	–0.30	<0.01
>74	–0.64	<0.01	–0.63	<0.01	–0.57	<0.01	–0.47	<0.01
Sex								
Male	0.06	0.40	0.06	0.42	0.05	0.44	0.01	0.96
Ethnicity/race								
White Hispanic	0.20	<0.01	0.21	<0.01	0.17	<0.01	0.01	0.99
Black Hispanic	0.12	0.52	0.15	0.44	0.18	0.34	0.06	0.90
Native American	0.19	0.27	0.22	0.21	0.18	0.27	0.17	0.52
African-American	0.29	<0.01	0.30	<0.01	0.25	<0.01	0.31	<0.01
Asian	0.54	0.02	0.54	0.02	0.35	<0.01	1.34	<0.01
Caucasian	—	—	—	—	—	—	—	—
Service connection rating (%)								
50–100	–0.13	<0.01	–0.14	<0.01	–0.12	<0.01	–0.03	0.54
<50	–0.08	0.01	–0.08	0.01	–0.07	0.03	–0.04	0.52
0	—	—	—	—	—	—	—	—
Care seeking (utilization) factors								
Pharmacy cost (log)	—	—	0.07	<0.01	–0.06	<0.01	–0.14	<0.01
Number of outpatient visits (log)	—	—	–0.06	<0.01	–0.07	<0.01	–0.14	<0.01
Geographic factor								
Metropolitan areas	—	—	–0.04	0.10	–0.04	0.11	–0.04	0.47
Diabetes duration/severity and comorbidity factors								
Amputation or ulceration	—	—	—	—	0.18	<0.01	0.22	<0.01
Days of hospitalization (log)	—	—	—	—	–0.01	0.38	–0.08	<0.01
Treatment modality								
Oral medication	—	—	—	—	1.85	<0.01	2.17	<0.01
Insulin	—	—	—	—	1.42	<0.01	1.56	<0.01
Oral and insulin therapies	—	—	—	—	2.49	<0.01	2.73	<0.01
Diet control only	—	—	—	—	—	—	—	—
Monitored blood glucose	—	—	—	—	0.23	<0.01	0.34	<0.01
Charlson index	—	—	—	—	—	—	0.09	<0.01

Prediction variables were entered stepwise with each additional model. Model 1 contained only demographic variables and Model 4 contained all of the predictors. This stepwise construction of risk adjustment enabled us to view potential confounding between predictive factors. For example, pharmacy cost was positively (0.07) associated with levels of HbA_{1c} shown in Model 2. That association, however, turned negative (–0.06) when diabetes duration/severity factors were introduced in Model 3. That negative association more than doubled (–0.14) when a Charlson index comorbidity score was introduced in Model 4, which suggested further disentangling of confounding among those variables.

risk HbA_{1c}) were responsive to risk adjustment. In this exploratory model, risk adjusting the level actually resulted in less change than for the high-risk rate, a finding that must be validated.

Recent efforts to use risk adjustment for HbA_{1c} levels have incorporated patient-level factors obtained from surveys (41). The feasibility of this approach remains in question, however. Ideally, a risk adjustment method for HbA_{1c} should be widely applicable and inexpensive to use. For this reason, we developed this exploratory model based solely on existing administrative data, acknowledging that such data are, out of necessity, often proxies for the desired variable. For example, duration of diabetes, a key patient-level factor closely correlated with glycemic control, is not available in administrative data; we used treatment modality as a proxy measure. (Of note, we also used treatment modality to control for potential sampling bias introduced by the choices in identifying diabetic patients permitted by the Health Plan Employer Data and Information Set [HEDIS].) Currently, health plans can use pharmacy records, ICD-9-CM codes, or a combination of the 2, possibly favoring plans with larger numbers of enrollees with diet-controlled diabetes.

Several results of our study reassured us of the validity of our approximations. First, the coefficient estimates for our constructed measures were in the expected directions. Our findings were consistent with previous studies of demographic variables and glycemic control (14,42) that confirmed that age and ethnicity were associated with HbA_{1c}. These findings

Appendix 2—*Predicting high-risk HbA_{1c} (>9.5%) using a generalized hierarchical mixed-effects model*

	Model 1		Model 2		Model 3		Model 4	
	Estimate	*P*	Estimate	*P*	Estimate	*P*	Estimate	*P*
Conditional Veterans Affairs average level	0.27	<0.01	0.25	<0.01	0.37	<0.01	0.06	<0.01
Demographic factors								
Age (years)								
<45	0.52	<0.01	0.52	<0.01	0.43	<0.01	0.60	<0.01
45–54	0.27	<0.01	0.28	<0.01	0.26	<0.01	0.33	<0.01
55–64	—	—	—	—	—	—	—	—
65–74	−0.38	<0.01	−0.38	<0.01	−0.37	<0.01	−0.33	<0.01
>74	−0.69	<0.01	−0.69	<0.01	−0.66	<0.01	−0.52	<0.01
Sex								
Male	0.12	0.17	0.12	0.18	0.12	0.19	0.14	0.40
Ethnicity/race								
White Hispanic	0.27	<0.01	0.28	<0.01	0.25	<0.01	0.14	0.32
Black Hispanic	0.22	0.30	0.24	0.28	0.28	0.21	0.03	0.95
Native American	0.26	0.19	0.28	0.17	0.24	0.24	0.32	0.29
African-American	0.28	<0.01	0.29	<0.01	0.25	<0.01	0.28	<0.01
Asian	0.52	0.04	0.52	0.04	0.60	0.02	0.88	0.06
Caucasian	—	—	—	—	—	—	—	—
Service connection rating (%)								
50–100	−0.16	<0.01	−0.17	<0.01	−0.14	<0.01	−0.04	0.53
<50	−0.13	<0.01	−0.13	<0.01	−0.12	<0.01	−0.09	0.23
0	—	—	—	—	—	—	—	—
Care seeking (utilization) factors								
Pharmacy cost (log)	—	—	0.03	0.02	−0.08	<0.01	−0.14	<0.01
Number of outpatient visits (log)	—	—	−0.03	0.13	−0.04	0.05	−0.11	<0.01
Geographic factor								
Metropolitan areas	—	—	−0.07	0.03	−0.07	0.04	−0.05	0.40
Diabetes duration/severity and cormorbidity factors								
Amputation or ulceration	—	—	—	—	0.18	0.01	0.22	<0.01
Days of hospitalization (log)	—	—	—	—	−0.00	0.92	−0.07	<0.01
Treatment modality								
Oral medication	—	—	—	—	2.45	<0.01	3.16	<0.01
Insulin	—	—	—	—	2.08	<0.01	2.62	<0.01
Oral and insulin therapies	—	—	—	—	3.06	<0.01	3.70	<0.01
Diet control only	—	—	—	—	—	—	—	—
Monitored blood glucose	—	—	—	—	0.18	<0.01	0.27	<0.01
Charlson index	—	—	—	—	—	—	0.06	0.03

See legend for APPENDIX 1.

alone justify risk adjustment of accountability measures. In addition, we observed sizable changes of effect estimates in the expected direction with entry into the model of disease severity and comorbidity adjusters, including the modality of treatment and the Charlson index. These changes should be distinguished from regression to the mean, because the latter occurs only in longitudinal settings. The inverse correlation of outpatient utilization with HbA_{1c} levels was understandable, albeit different from a recent analysis of the NHANES 3 (14), when viewed in the context of multiple risk adjustments with a carefully specified statistical model. Although preliminary, these findings suggest that the associations between health-seeking/utilization factors and HbA_{1c} were likely to be confounded by severity and diabetes duration as well as inpatient comorbidity, underscoring the importance of controlling for all of these variables to estimate their respective effects.

We should stress that this study focused on the effects of applying risk adjustment to a selected group of health care systems and not on describing diabetes care in the VHA overall. To facilitate our analysis, we chose to exclude individuals without complete data and recognized the potential selection bias this introduced. However, the number of facilities reporting was reasonably large and the variability in HbA_{1c} was broad, thereby making possible an exploration of how risk adjustment might affect a quality assessment.

We also note that only 65 of 173 facilities chose to report to NCCC in 1996, a situation reminiscent of voluntary HEDIS reporting, in which about half of health plans still choose not to participate. Ongoing mandated data collection in the VHA should allow future analyses of a more sizable proportion of the population, and the issue of missing data will then need to be addressed accordingly.

Although our ability to obtain cross-linked demographic, laboratory, pharmacy, and diagnostic coding data from the VHA database may not be widely possible to duplicate in the health care industry currently, availability of such data is anticipated in the future. Even though our veteran population was predominantly male, diabetes severity and outcomes have not previously been demonstrated to be related to sex (14); thus, our findings may well apply to other settings that include more women.

Several important variables that could potentially influence HbA_{1c} were not available as direct measures. We lacked data on socioeconomic and educational status, but these factors have not been correlated with glycemic control in non-Hispanic white or Mexican-American (43), Caucasian (44), or African-American (44,45) patients.

Comorbidity may impact diabetes control. Because of our particular data characteristics, we were unable to use non–disease-specific risk adjustment developed for use with outpatient ICD-9-CM codes (such as the Adjusted Care Groups [16,46]) or in-depth pharmacy information (such as the Chronic Disease Score [47]). We did have inpatient ICD-9-CM codes and therefore applied an adapted Charlson index to estimate comorbidities, albeit only for the minority of our sample who were hospitalized during FY96 (21). Further validation using outpatient ICD-9-CM data is warranted.

Medical use and expenditure patterns have not been captured completely in this study. Again because of data limitations, we were unable to build on previous findings of outpatient ICD-9-CM codes correlating with medical use and expenditure patterns for Medicare beneficiaries with diabetes (48,49). Data limitations also prevented us from fully describing resource utilization, because many veterans are also Medicare beneficiaries and receive at least some private-sector care (50).

A crucial concern in observational studies of this nature is the inability to ascribe directionality and causation. For example, pharmacy costs and the number of outpatient contacts were found to be significantly associated with high-risk HbA_{1c} rates in our study. Based on this work, we cannot distinguish whether more frequent contacts or higher pharmacy spending are a result of more severe disease or of better systematic health care delivery. Ongoing data collection should allow longitudinal analyses to better explore these relationships.

Summary

We demonstrated the feasibility of evaluating quality of care of medical providers using administrative data sources. We developed an exploratory diabetes-specific hierarchical mixed-effects model to predict HbA_{1c}. Performance status of diabetes care measured by HbA_{1c} levels or rates changed for a substantial number of facilities, especially with the current DQIP accountability measure for HbA_{1c} (percentage of patients with $HbA_{1c} > 9.5\%$). Although our model remains to be validated in other databases, we demonstrated that use of non–risk-adjusted HbA_{1c} values may inappropriately identify plans as high or low performers.

Acknowledgments — This study was supported by a grant from Veterans Integrated Service Network for New York-New Jersey and the Veterans Affairs Health Services Research and Development Service (Quality Enhancement Research Initiative).

References

1. Herman WH, Eastman RC: The effects of treatment on the direct costs of diabetes. *Diabetes Care* 21 (Suppl. 3):C19–C24, 1998
2. U.K. Prospective Diabetes Study Group: Intensive blood-glucose control with sulphonylureas or insulin compared with conventional treatment and risk of complications in patients with type 2 diabetes (UKPDS 33). *Lancet* 352:837–853, 1998
3. Clark CM, Lee DA: Prevention and treatment of the complications of diabetes mellitus. *N Engl J Med* 332:1210–1217, 1995
4. Testa MA, Simonson DC: Health economic benefits and quality of life during improved glycemic control in patients with type 2 diabetes mellitus. *JAMA* 280:1490–1496, 1998
5. Diabetes Quality Improvement Project Initial Measure Set (Final Version): 14 August 1998. http://www.diabetes.org/
6. Donabedian A: Basic approaches to assessment: structure, process, and outcome. In *The Definition of Quality and Approaches to Its Assessment*. Ann Arbor, MI, Health Administration Press, 1980, p. 77–128
7. Palmer RH: Quality of care. *JAMA* 277: 1896–1897, 1997
8. Chassin MR, Galvin RW, the National Roundtable on Health Care Quality: The urgent need to improve health care quality. *JAMA* 280:1000–1005, 1998
9. Mitka M: Renewed efforts to improve the quality of health care. *JAMA* 281:404, 1999
10. Pogach LM, Hawley G, Weinstock R, Sawin C, Schiebe H, Cutler F, Zieve F, Bates M, Repke D: Diabetes prevalence and hospital and pharmacy use in the Veterans Health Administration (1994): use of an ambulatory care pharmacy-derived database. *Diabetes Care* 21:368–373, 1998
11. Naito H, Hartfiel J, Little R, Pogach L: Standardization of glycohemoglobin (GHB) methods in the Department of Veterans Affairs Medical Centers. American Diabetes Association Annual Meeting, San Francisco, CA, June 1996
12. Naito H, Hartfiel J, Eckfeldt J, Pogach L: Matrix-induced biases observed with proficiency testing survey materials that are used for glycohemoglobin testing. American Diabetes Association Annual Meeting, San Francisco, CA, June 1996
13. Chassin MR, Hannan EL, DeBuono BA: Benefits and hazards of reporting medical outcomes publicly. *N Engl J Med* 334:394–398, 1996
14. Harris MI, Eastman R, Cowie CC, Flegal KM, Eberhardt MS: Racial and ethnic differences in glycemic control of adults with type 2 diabetes. *Diabetes Care* 22:403–408, 1999
15. Harris MI, Flegal KM, Cowie CC, Eberhardt MS, Goldstein DE, Little RR, Weidmeyer HM, Byrd-Holt DD: Prevalence of diabetes, impaired fasting glucose tolerance in US adults: the Third National Health and Nutritional Examination Survey, 1988–1994. *Diabetes Care* 21:518–524, 1998
16. Weiner JP, Parente ST, Garnick DW, Fowles J, Lawthers AG, Palmer H: Variation in office-based quality: a claims-based profile of care provided to Medicare patients with diabetes. *JAMA* 273:1503–1508, 1995
17. Brook RH, Kamberg CJ, McGlynn EA: Health system reform and quality. *JAMA* 276:476–480, 1996
18. Keeler EB, Rubenstein LV, Kahn KL, Draper D, Harrison ER, McGinty MJ, Rogers WH, Brook RH: Hospital characteristics and quality of care. *JAMA* 268:1709–1714, 1992
19. Kenny SJ, Aubert RE, Geiss LS: Prevalence and incidence of non-insulin-dependent diabetes. In *Diabetes in America*, 2nd ed. Harris MI, Cowie CC, Stern MP, Boyko EJ, Reiber GE, Bennett PH, Eds. Washington, DC, U.S. Govt. Printing Office, 1995, p. 47–67 (NIH publ. no. 95-1468)
20. Charlson ME, Pompei R, Ales KL, MacKenzie CR: A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis* 40:373–383, 1987
21. Deyo RA, Cherkin DC, Ciol MA: Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol* 45:613–619, 1992
22. Bryk AS, Raudenbush SW: *Hierarchical Linear Models: Applications and Data Analysis Methods*. Newbury Park, CA, Sage Publications, 1992
23. Longford NT: *Random Coefficient Models*. New York, Oxford University Press, 1993
24. Hu FB, Goldberg J, Hedeker D, Henderson

WG: Modelling ordinal responses from co-twin control studies. *Stat Med* 17:957–970, 1998
25. Akaike H: A new look at the statistical model identification. *IEEE Trans Autom Control* 19:716–723, 1974
26. McBean AM, Hebert P, Engelgau M, Geiss L, Tierney E: *National Diabetes Cohort*. Tampa, FL, Centers for Disease Control and Prevention Translation Conference, 1998
27. American Diabetes Association: *Economic Consequences of Diabetes Mellitus in the U.S. in 1997*. Alexandria, VA, American Diabetes Association, 1998, p. 296–309
28. Selby JV, Ray GT, Zhang D, Coby CJ: Excess costs of medical care for patients with diabetes in a managed care population. *Diabetes Care* 20:1396–1402, 1997
29. Javitt JC, Chiang Y: Economic impact of diabetes. In *Diabetes in America*, 2nd ed. Harris MI, Cowie CC, Stern MP, Boyko EJ, Reiber GE, Bennett PH, Eds. Washington, DC, U.S. Govt. Printing Office, 1995, p. 601–612 (NIH publ. no. 95-1468)
30. Eastman R, Javitt JC, Herman W, Dasbach EJ, Zbrozek AS, Dong F, Manninen D, Garfield SA, Copley-Merriman C, Maier W, Eastman JF, Kotsanos J, Cowie CC, Harris M: Model of complications of NIDDM: model construction and assumptions. *Diabetes Care* 20:725–734, 1997
31. Harris MI, Eastman RC, Siebert C: The DCCT and medical care for diabetes in the U.S. *Diabetes Care* 17:761–764, 1994
32. Fertig BJ, Simmons DA, Martin DB: Therapy for diabetes. In *Diabetes in America*, 2nd ed. Harris MI, Cowie CC, Stern MP, Boyko EJ, Reiber GE, Bennett PH, Eds. Washington, DC, U.S. Govt. Printing Office, 1995, p. 519–540 (NIH publ. no. 95-1468)
33. DCCT Group: Lifetime benefits and costs of intensive therapy as practices in the Diabetes Control and Complications Trial. *JAMA* 276:1409–1415, 1996
34. U.K. Prospective Diabetes Study 34: Effect of intensive blood-glucose control with metformin on complication in overweight patients with type 2 diabetes. *Lancet* 352: 854–865, 1998
35. U.K. Prospective Diabetes Study 40: Cost effectiveness analysis of improved blood pressure control in hypertensive patients with type 2 diabetes. *BMJ* 317:720–726, 1998
36. Brinkley J: U.S. releasing lists of hospitals with abnormal mortality rates. *New York Times*. 12 March 1986: Sect. A, p. 1
37. Daley J, Jencks S, Draper D, Lenhart G, Thomas N, Walker J: Predicting hospital-associated mortality for Medicare patients: a method for patients with stroke, pneumonia, acute myocardial infarction and congestive heart failure. *JAMA* 260:3617–3624, 1988
38. Iezzoni LI, Shwartz M, Ash AS, Hughes JS, Kaley J, Mackiernan YD: Using severity-adjusted stroke mortality rates to judge hospitals. *Int J Qual Health Care* 7:81–94, 1995
39. Hannan EL, Siu AL, Kumar D, Kilburn H Jr., Chassin MR: The decline in coronary artery bypass graft surgery mortality in New York State: the role of surgeon volume. *JAMA* 273:209–213, 1995
40. Jollis JG, Romano PS: Pennsylvania's focus on heart attack: grading the score card. *N Engl J Med* 338:983–987, 1998
41. 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* 281:2098–2105, 1999
42. Tull ES, Roseman JM: Diabetes in African Americans. In *Diabetes in America*, 2nd ed. Harris MI, Cowie CC, Stern MP, Boyko EJ, Reiber GE, Bennett PH, Eds. Washington, DC, U.S. Govt. Printing Office, 1995, p. 613–630 (NIH publ. no. 95-1468)
43. Haffner SM, Hazuda HP, Stern MP, Patterson JK, Van Heuven WAJ, Fong D: Effect of socioeconomic status on hyperglycemia and retinopathy levels in Mexican Americans with NIDDM. *Diabetes Care* 12:128–134, 1989
44. Blaum CS, Velez L, Hiss RG, Halter JB: Characteristics related to poor glycemic control in NIDDM patients in community practice. *Diabetes Care* 20:7–11, 1997
45. Eberhardt MS, Lackland DT, Wheeler FC, German RR, Teuttsch SM: Is race related to glycemic control? An assessment of glycosylated hemoglobin in two South Carolina communities. *J Clin Epidemiol* 47:1181–1189, 1994
46. Iezzoni LI: Risk and outcomes. In *Risk Adjustment for Measuring Healthcare Outcomes*. Iezzoni LI, Ed. Chicago, Health Administration Press, 1997, p. 1–42
47. Von Korff M, Wagner EH, Saunders K: A chronic disease score from automated pharmacy data. *J Clin Epidemiol* 45:197–203, 1992
48. Krop JS, Powe NR, Weller WE, Shaffer TJ, Saudek CD, Anderson GF: Patterns of expenditures and use of services among older adults with diabetes: implications for the transition to capitated managed care. *Diabetes Care* 21:747–752, 1998
49. *Medicare: Most beneficiaries with diabetes do not receive recommended monitoring services*. General Accounting Office, Department of Health and Human Services, Report 97-48, 28 March 1997
50. Fleming C, Fisher ES, Chang CH, Bubolz TA, Malenka DJ: Studying outcomes and hospital utilization in the elderly: the advantages of a merged data base for Medicare and Veterans Affairs hospitals. *Med Care* 30:377–391, 1992