

Estimating Physician Effects on Glycemic Control in the Treatment of Diabetes

Methods, effects sizes, and implications for treatment policy

PETER W. TUERK, PHD^{1,2}
MARTINA MUELLER, PHD^{3,4}
LEONARD E. EGEDE, MD, MS^{4,5}

OBJECTIVE — Researchers have only just begun to investigate physician-related effects on medical outcomes. Such research is necessary for developing empirically informed practice guidelines and policy. The primary goal of this study was to investigate whether glucose management in type 2 diabetes varies by randomly assigned physicians over the course of a year in treatment. A second goal of the study was to investigate whether physician-related effects vary across differential patient characteristics. A tertiary goal was to investigate potential patient-level effects on glucose management.

RESEARCH DESIGN AND METHODS — Hierarchical linear models were used to investigate A1C among 1,381 patients, nested within 42 randomly assigned primary care physicians at a Veterans Affairs medical center in the southeastern U.S. The primary outcome measure was change in A1C over the course of 1 year in treatment. On average, each study physician had 33 patients with diabetes.

RESULTS — Overall, physician-related factors were associated with statistically significant but modest variability in A1C change (2%), whereas patient-level factors accounted for the majority of variation in A1C change (98%). Physician effects varied by patient characteristics, mattering more for black patients, patients aged 65 years, and patients whose glucose management improved over the treatment year.

CONCLUSIONS — The results of this study indicate that differential physician effects have minimal impact on glycemic control. Results suggest that it is logical to support policies encouraging the development of patient-level behavioral interventions because that is the level that accounts for the majority of variance in glycemic control.

Diabetes Care 31:869–873, 2008

Diabetes affects ~20.8 million people in the U.S. or 7% of the population (1). Diabetes is associated with mortality, morbidity, increased health care utilization, and increased health care costs (1). Although comprehensive treatment can prevent or delay complications from diabetes (2–5), data

indicate that glycemic control in patient populations is poor (6) and that diabetes care in the U.S. is often suboptimal (7).

Both patient and physician factors play roles in the successful treatment of diabetes. Recent studies indicate that a substantial proportion of diabetes management (e.g., diet, exercise, glucometer

use, and medications) is under the direct control of patients (8); however, the majority of medical care and education for diabetes is provided by primary care physicians (9,10). Thus, physician factors may substantially contribute to patient behavioral change and outcomes. The health care industry currently spends substantial resources tracking physician performance measures to address health care deficiencies and promote accountability (11–13). Consistent with this agenda, “pay-for-performance” policies are gaining popularity among insurance providers to encourage physicians to meet behavioral targets and avoid errors (14,15). Although it makes good sense to track administrative and procedural performance measures and to consider alternative ways of providing motivation to physicians, doing so is expensive and has not been conclusively shown to improve outcomes or change physician behavior (16).

There are few studies that have examined physician effects in diabetes outcomes (16). To address this gap in the literature, the current study used Veterans Health Administration data and hierarchical linear modeling (HLM) to investigate glucose management as measured by A1C change over the course of 1 year in treatment. The primary goal of the study was to measure physician-related effects on A1C change while controlling for patient-level factors. A second goal of the study was to investigate whether physician-related effects on A1C varied across patient characteristics (i.e., race, age, and health status). A tertiary goal of the study was to provide estimates for patient-level effects on A1C outcomes.

RESEARCH DESIGN AND METHODS

— We created a cohort of patients with type 2 diabetes at a Veterans Affairs medical center in the southeastern U.S. using multiple patient files and existing Veterans Health Administration data sources. Individuals with type 2 diabetes were identified based on having at least two ICD-9 codes for diabetes and having two or more visits each year since their diagnosis based on a previously validated

From the ¹Department of Psychiatry and Behavioral Sciences, Medical University of South Carolina, Charleston, South Carolina; the ²Psychology Service, Ralph H. Johnson Veterans Affairs Medical Center, Charleston, South Carolina; the ³Department of Biostatistics, Bioinformatics, and Epidemiology, Medical University of South Carolina, Charleston, South Carolina; the ⁴Department of Medicine, Center for Health Disparities Research, Medical University of South Carolina, Charleston, South Carolina; and the ⁵Veterans Affairs Targeted Research Enhancement Program (TREP), Ralph H. Johnson Veterans Affairs Medical Center, Charleston, South Carolina.

Corresponding author: Peter Tuerk, PhD, MUSC 165 Cannon St., Third Fl., P.O. Box 250852, Charleston, SC 29425. E-mail: tuerk@musc.edu.

Received for publication 23 August 2007 and accepted in revised form 30 January 2008.

Published ahead of print at <http://care.diabetesjournals.org> on 19 February 2008. DOI: 10.2337/dc07-1662.

Abbreviations: HLM, hierarchical linear modeling; ICC, intraclass correlation; RCT, randomized controlled trial.

© 2008 by the American Diabetes Association.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

algorithm (17). The cohort included 14,931 patients. For this study, we selected patients from the cohort who had at least two A1C measurements within 1 year apart and who had a specific designated primary care provider within the same period. To ensure reliability of study estimates, we excluded patients whose providers had less than five diabetic patients. The resulting dataset consisted of 1,381 patients nested within 42 physicians. On average, each study physician had (mean \pm SD) 33 ± 12.99 patients with diabetes.

Demographics

Patients in the study sample were 98% male. The mean age \pm SD of patients in the sample was 66 ± 11 years and the baseline A1C was 7.52 ± 1.97 . Race and ethnicity measures were aggregated to create three racial/ethnic groups: non-Hispanic whites (whites, 52%), non-Hispanic blacks (blacks, 38%), and Hispanics, Asians, multi-racial, and unknown (other, 10%).

Demographics by physician specification

Preliminary HLM analyses were used to investigate the distributions of patient characteristics among study physicians to assure random distribution and identify potential covariates. There were no statistically significant differences among physicians in regards to patient sex, minority status, systolic blood pressure, diastolic blood pressure, or weight. Statistically significant, though small, differences were found in baseline A1C ($<1\%$ of the variance) and percentage of patients aged 65 years (3% of the variance). Thus, both were identified as covariates in the full model.

Primary outcome measure

The outcome measure of interest in the current study was change in A1C from baseline. A1C measures the average percentage of blood glucose over the prior 2–3 months and is considered an index of how well glucose is being controlled. Complications from diabetes can be prevented or delayed if A1C levels are kept at $<7\%$ (18). For the current study, A1C was measured at baseline and again at an average of 12 months after baseline (mean \pm SD 363 ± 36 days).

Statistical analyses

HLM was used to address the current research questions because it allows variance in outcomes to be partitioned

between patients and physicians. This is possible due to the method of handling of error. In regular multiple regression models, prediction error associated with patient characteristics and prediction error associated with physician specification would be combined into one error term. In contrast, HLM allows each physician to have a unique intercept (and slope) for patient effects on outcomes. Thus, error attributed to patient effects and error attributed to physician effects are in different terms (19). Statistical analyses in the current study were performed using HLM6 software (20). Veterans Affairs data are well suited for such an analytic strategy because patients are randomly assigned to available physicians. This limits confounding factors by naturally encouraging (though not guaranteeing) an orthogonal relationship between patient and physician characteristics.

Physician effects

To investigate the role of physician effects on A1C change over and above patient-level factors, a hierarchical unconditional model was used. In HLM, such models are called unconditional because they do not include specific predictors. Their function is to partition variance in outcomes among level 1 and level 2 unit specifications (21). An unconditional model was used to test the null hypothesis that there was no variation in A1C due to physician-level factors. To do this, variance components σ^2 and τ were estimated for patient-level and physician-level error terms, respectively. The null hypothesis could be safely rejected if variance component τ was significantly different from zero and significantly improved model fit. Significance for model fit was based on comparing pairs of model deviances distributed asymptotically on a χ^2 distribution (20).

Given significant results, unconditional model effect sizes were estimated using intraclass correlations (ICCs) consistent with the methods recommended by Bryk and Raudenbush (21). In a two-level hierarchical structure, the ICC is defined as the proportion of the total variance in the outcome variable that is between the level 2 units, or R^2_{total} (19). In a hierarchical patient-physician model, the ICC represents the proportion of variance accounted for at the physician level. It therefore follows that the remainder of the variance can be attributed to factors at the patient level or to other unspecified nested factors.

Physician effects by patient characteristics

To investigate if and how physician effects varied across patient populations, a simple effects design was used. Separate unconditional models were estimated for blacks, whites, others, patients aged 65 years, patients aged <65 years, and patients who improved, stayed the same, or got worse over the course of the treatment year. Patient improvement was defined as a $\geq 15\%$ decrease in A1C from baseline (in the study sample, this averaged to a 1-point overall decrease in A1C). Patient worsening was defined as a $\geq 15\%$ increase in A1C, and patients who stayed the same were defined as those whose A1C change was $<15\%$ in either direction.

Patient effects

To investigate the role of available patient characteristics in diabetes outcomes, a stepwise random coefficients model was estimated. The model was constructed by adding specific patient-level predictors to the total-sample unconditional model. Predictors were entered in the following order: baseline A1C, age, and race/ethnicity (blacks, whites, or others), with each race/ethnic distinction dummy coded and alternately added for effect estimation. With the current sample size ($n = 1,381$), it is assumed that coefficient estimates for patient age and race are reliable and valid for similar populations; however, mathematical coupling and regression to the mean pose potential threats to the precision of the correlation between baseline A1C and change in A1C (22). Accordingly, baseline A1C is included as a peripheral control rather than as a definitive measure of the impact of initial disease on A1C change.

As mentioned, within an HLM framework the ICC or R^2_{total} indicates the amount of total variance accounted for by level 1 and level 2 factors, regardless of what those factors are. By adding specific predictors to each level, it becomes possible to estimate R^2_{within} and R^2_{between} . In the present study, R^2_{within} was estimated for each level 1 predictor in a manner consistent with the recommendations of Snijders and Bosker (22) and denoted what portion of the total variance attributed to patients could be accounted for by each patient predictor. R^2_{between} was not estimated because no specific physician-level factors or characteristics were entered into the models.

Table 1—Variance components for unconditional models: physician effects on diabetes A1C outcomes

| | Random effect | Variance component* | d.f. | χ^2 | P | ICC (R^2_{total}) |
|----------------------------|---------------------|---------------------|-------|----------|--------|------------------------------|
| Overall sample | Patient: σ^2 | 5.14 | 1,381 | | | |
| | Physician: τ | 0.09 | 41 | 71.77 | 0.002 | 0.02 |
| Black | Patient: σ^2 | 6.57 | 528 | | | |
| | Physician: τ | 0.18 | 41 | 62.22 | 0.018 | 0.03 |
| White | Patient: σ^2 | 3.76 | 719 | | | |
| | Physician: τ | 0.08 | 41 | 55.52 | 0.064 | NS |
| Aged >65 years | Patient: σ^2 | 3.97 | 718 | | | |
| | Physician: τ | 0.13 | 41 | 60.64 | 0.024 | 0.02 |
| Aged <65 years | Patient: σ^2 | 6.04 | 663 | | | |
| | Physician: τ | 0.01 | 41 | 47.54 | 0.224 | NS |
| Patients who improved† | Patient: σ^2 | 2.74 | 461 | | | |
| | Physician: τ | 0.14 | 41 | 68.20 | 0.005 | 0.05 |
| Patients who remained same | Patient: σ^2 | 0.87 | 512 | | | |
| | Physician: τ | 0.01 | 41 | 48.37 | 0.17 | NS |
| Patients who got worse | Patient: σ^2 | 1.51 | 408 | | | |
| | Physician: τ | 0.00 | 40 | 28.25 | >0.500 | NS |

*Variance components σ^2 and τ denote an estimate of the amount of variance in outcomes attributable to patient and physician effects, respectively. While patient effects are assumed to account for variance in A1C outcomes, physician effects were tested with χ^2 to compare the fit of models that included physician effects with null models without physician effects. A significant χ^2 indicates rejection of the null hypothesis that patient outcomes do not vary by randomly assigned physicians. †Patients who improved were defined as patients whose A1C levels improved by $\geq 15\%$ of baseline, patients who remained the same were those whose levels stayed within 15% of baseline, and patients who got worse were those whose A1C levels rose by $> 15\%$ of baseline. NS, not significant.

RESULTS

Physician effects

The mean \pm SD A1C after a year in treatment was 7.21 ± 1.68 . Table 1 shows the estimated variance components ($\sigma^2 = 5.14$; $\tau = 0.09$) for patient- and physician-level units in the overall-sample unconditional model. The value of τ was significantly different from zero and improved model fit ($\chi^2 = 71.77$; d.f. = 41; $P = 0.002$), indicating the presence of physician-level effects beyond patient-level factors. The ICC was 0.02, indicating that 2% of the overall variance in A1C change over the study period could be accounted for by physician-level factors and 98% could be attributed to patient-level or unspecified factors. As Table 2 indicates, adding patient covariates of age, baseline A1C, and nonminority status to the unconditional model had no substantial effect on the ratio of variance accounted for by physician specification ($\Delta\text{ICC} = 0.005$). Figure 1 shows the amount of variance in A1C change accounted for by patient and physician factors.

Physician effects by race/ethnicity. The unconditional model for black patients yielded a physician-level variance component ($\tau = 0.18$) that was significantly different from zero and improved

model fit ($\chi^2 = 62.22$; d.f. = 41; $P = 0.02$), indicating the presence of physician-level effects beyond patient-level factors. The ICC indicated that 3% of the overall variance in A1C change for blacks could be accounted for by physician-level factors. The unconditional model for whites yielded a physician-level variance component ($\tau = 0.08$) that was not significantly different from zero and did not improve model fit ($\chi^2 = 55.52$; d.f. = 41; $P = 0.06$). However, marginally significant findings for whites may have been due to a limited sample size rather than to a true difference between groups. The nonsignificant ICC indicated that 2% of

the overall variance in A1C change for whites could be accounted for by physician-level factors.

Physician effects by age. The unconditional model for patients aged 65 years yielded a physician-level variance component ($\tau = 0.13$) that was significantly different from zero and improved model fit ($\chi^2 = 60.64$; d.f. = 41; $P = 0.02$). The ICC was 0.02, indicating that 2% of the variance in A1C change for patients aged 65 years could be accounted for by physician-level factors. The unconditional model for patients aged < 65 years yielded a physician-level variance component ($\tau = 0.01$) that was not significantly different from zero ($\chi^2 = 47.54$; d.f. = 41; $P = 0.22$).

Physician effects by change in A1C. The unconditional model for patients who improved yielded a physician-level variance component ($\tau = 0.14$) that was significantly different from zero and improved model fit ($\chi^2 = 68.20$; d.f. = 41; $P = 0.005$). The ICC was 0.05, indicating that 5% of the overall variance in A1C for patients who improved could be accounted for by physician-level factors. The unconditional model for patients who did not improve yielded a physician-level variance component ($\tau = 0.01$) that was not significantly different from zero and hence did not improve model fit ($\chi^2 = 48.37$; d.f. = 41; $P = 0.17$). The unconditional model for patients who got worse yielded a physician-level variance component ($\tau = 0.00$) that was not significantly different from zero and hence did not improve model fit ($\chi^2 = 28.25$; d.f. = 40; $P > 0.50$).

Total sample patient effects

Baseline A1C. The random coefficients model predicting A1C change estimated effects for patient A1C baseline, patient age, and patient race. The first factor entered into the model was baseline A1C. As

Table 2—Random coefficients model predicting A1C change

| | Coefficient | SEM | T-ratio | P | R^2_{within} |
|-------------------------|--------------------|-------|----------|-------|-----------------------|
| Fixed effect | | | | | |
| Intercept | 6.08 | 0.33 | 18.23 | 0.000 | |
| A1C baseline | −0.73 | 0.02 | −31.94 | 0.000 | 0.46 |
| Patient age | −0.01 | 0.004 | −2.27 | 0.023 | 0.004 |
| Nonminority | −0.21 | 0.09 | −2.33 | 0.020 | 0.003 |
| Random effect | | | | | |
| | Variance component | d.f. | χ^2 | | |
| A1C intercept (u_0) | 0.034 | 41 | 58.08 | 0.040 | |
| Level 1 (r) | 2.76 | | | | |

d.f. within = 1,380; d.f. between = 41.

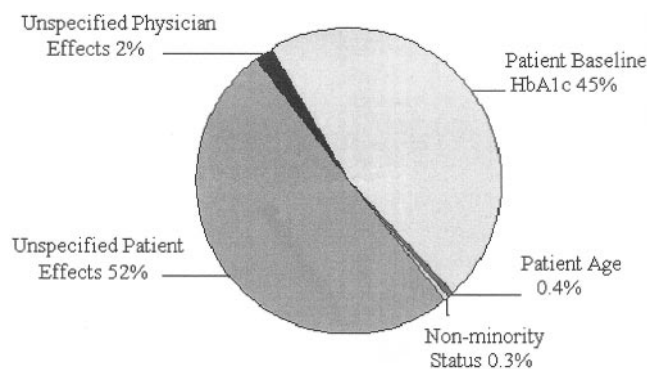


Figure 1—Total variance accounted for in A1C change by patient and physician factors.

Table 2 indicates, baseline measurement significantly predicted A1C improvement over the course of the treatment year ($t = -31.94$; d.f. = 1,380; $P = 0.000$). Higher baselines were associated with greater decreases in A1C. The estimated R^2_{within} was 0.46, indicating that baseline levels accounted for 46% of the patient-level variance in A1C change or 45% of the total variance (46 of 98%).

Patient age. Patient age significantly predicted A1C change over 12 months ($t = -2.27$; d.f. = 1,380; $P = 0.023$) after controlling for baseline A1C. Older age was associated with greater A1C improvement. The estimated R^2_{within} was 0.004, indicating that patient age accounted for 1% of patient-level variance in A1C change. Even so, the effect translated into measurable differences in health outcomes. A production function of the modeled age coefficient indicated that for a patient aged 55 years, age was associated with an improvement of 0.55 in A1C levels, whereas for a patient aged 75 years, the improvement was 0.75. This is an effect size of about one-tenth of an SD difference ($d = 0.09$).

Patient race/ethnicity. The three racial/ethnic groups were alternately entered into the model. Controlling for baseline A1C and age, categorization as black was not associated with significant change in A1C ($t = 0.15$; d.f. = 1,380; $P = 0.11$) and neither was categorization as other ($t = 0.04$; d.f. = 1,380; $P = 0.70$). However, categorization as white was associated with significant improvement in A1C over the course of the treatment year ($t = -2.33$; d.f. = 1,380; $P = 0.02$). Because the predictor variable was dummy coded, indicating being white versus black or other, the effect is referred to here as non-minority status. The estimated R^2_{within} was 0.003, indicating that nonminority status accounted for <1% of the patient-

level variance in A1C change. The effect translated into a between-group difference of about one-tenth of an SD ($d = 0.09$). Although in general black patients showed greater improvement than white patients, that difference appeared to be due to black patients having higher baseline A1C. After baseline A1C was entered into the model, white patients showed greater improvement.

CONCLUSIONS— The primary goal of this study was to investigate whether diabetes outcomes varied by randomly assigned physicians over the course of a year in treatment. Physician-level factors were associated with statistically significant but modest variability in health outcomes, accounting for 2% of the total variance in A1C change. Simple effects analyses revealed that physician factors mattered more for black than for white patients and more for patients aged >65 years than for those aged <65 years. Interestingly, physician factors mattered most for patients who improved over the year, accounting for 5% of the total variance in A1C change. Physician-level factors did not matter at all for patients who got worse. Significant patient-level predictors of A1C change were baseline A1C, age, and nonminority status. The results of this study suggest that physician effects only have minimal impact on glycemic control. These findings are consistent with the results of a prior study by Krein and colleagues (16). The researchers investigated 12,110 patients within 258 primary care physicians across 13 Veterans Affairs medical centers and found that physician effects accounted for 1% of total variance in A1C. The current study builds on those findings by examining change over time in A1C and by investigating how physician effects vary by patient characteristics.

The significance and meaning of the

physician-level effect reported here (2% of the variance) depend on one's perspective. Interpreting small effect sizes in large populations is often difficult. For example, in the U.S. population, Down's syndrome accounts for a statistically significant, though very small, amount of variance in IQ scores; however, one cannot assume Down's syndrome is not an important factor in IQ. Likewise, although physician effects appear to play a relatively small role in overall A1C outcomes for some individual patients or groups of patients, physician effects may be particularly important for survival and quality of life.

Alternatively, from a policy-oriented perspective, 2% of the variance does not seem a particularly fruitful target on which to focus limited resources and expensive interventions. Ninety-eight percent of the variance in A1C change in the current sample appears to be due to patient-level factors. Much of that variance is accounted for by largely immutable factors such as baseline disease, age, and race. However, a good deal of the variance (52%) remains to be explored and may very well be associated with potentially mutable factors.

The results of this study suggest it is logical to support policies encouraging the development of patient-level interventions because that level accounts for the majority of variance in glycemic control. Moreover, significant progress has been made in the development of effective patient-oriented interventions. A systematic review of 72 randomized controlled trials (RCTs) in patients with type 2 diabetes showed that self-management training improved diabetes knowledge, frequency and accuracy of glucose self-monitoring, and glycemic control (24). A meta-analysis of 31 RCTs found that self-management education decreased A1C on average by 0.76% (95% CI 0.34–1.18), and the amount of decrease was highly correlated to the number of contact hours (25). Moreover, a recent meta-analysis of 25 RCTs investigating the effects of psychological interventions on A1C outcomes indicated improved outcomes for the experimental conditions in about one-half of the studies (26). It should be noted that these effects were in addition to physician care. In essence, studies that found significant results did so by manipulating some factor within the “Unspecified Patient Effects” portion of Fig. 1. Given the current state of medical treatments, that portion is where the larg-

est potential payoff for future interventions exists.

Results of the current study are particularly relevant in light of the recent movement toward patient empowerment in diabetes self-care (27–28). As noted, results indicated that physician factors mattered the most for patients who improved and not at all for patients who got worse. The patient-empowerment literature could offer one straightforward explanation for this differential effect—patients who feel empowered may engage with their doctors more overall and, consequently, their doctors' skills, talents, and characteristics have more of an opportunity to affect outcomes. In contrast, patients who do not feel empowered to manage their diabetes may be less engaged with their physicians and thus the skills and characteristics of their physicians may not influence treatment outcomes. In such cases, "bad" treatment does not make patients worse and "good" treatment does not make them better. However, for patients who improved, physician-related factors made a meaningful difference in how much or how little they improved.

This study has some limitations. First, it is important to note that physician-level factors are not synonymous with physician characteristics. Potential nonrandom factors associated with physician specification, such as nursing team allocation or the availability of facility resources, may have the potential to bias estimates of physician-specific effects. Future investigations should take into account the ecology of hospital settings and sample from a variety of institutions. Second, there are important patient-level factors that were not available for this analysis. There is a need to further explore patient-level factors affecting diabetes outcomes, and future studies would benefit from including a broader range of patient-level predictors.

In conclusion, the results of this study suggest that it is logical to support policies encouraging the development of patient-level interventions because that level accounts for the majority of variance in glycemic control. Overall, results indicate a need for creativity in current treatment paradigms and policies, including the expanded use of multidisciplinary provider teams and behavioral approaches in the treatment of type 2 diabetes.

Acknowledgments—We thank Anouk Grubaugh, PhD, and Ron Acierno, PhD, Medical

University of South Carolina, and Sim Galazka, MD, University of Virginia School of Medicine, for helpful reviews of the manuscript.

References

1. National Institute of Health: National diabetes statistics [article online], 2005. Available from <http://diabetes.niddk.nih.gov/dm/pubs/statistics/index.htm#7>. Accessed 29 April 2007
2. Diabetes Control and Complications Trial Research Group: The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. *N Engl J Med* 329:977–986, 1993
3. UK Prospective Diabetes Study (UKPDS) 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
4. UK Prospective Diabetes Study (UKPDS) Group: Effect of intensive blood-glucose control with metformin on complications in overweight patients with type 2 diabetes (UKPDS 34). *Lancet* 352:854–865, 1998
5. UK Prospective Diabetes Study (UKPDS) Group: Tight blood pressure control and risk of macrovascular and microvascular complications in type 2 diabetes (UKPDS 38). *BMJ* 317:703–713, 1998
6. Harris MI, Eastman RC, 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
7. Saaddine JB, Engelgau MM, Beckles GL, Gregg EW, Thompson TJ, Narayan KM: A diabetes report card for the United States: quality of care in the 1990s. *Ann Intern Med* 136:565–574, 2002
8. Bodenheimer T, Wagner EH, Grumbach K: Improving primary care for patients with chronic illness: the chronic care model. II. *JAMA* 288:1909–1914, 2002
9. Egede LE: Lifestyle modification to improve blood pressure control in individuals with Diabetes: is physician advice effective? *Diabetes Care* 26:602–607, 2003
10. Harris MI: Medical care for patients with diabetes: epidemiologic aspects. *Ann Intern Med* 124:117–122, 1996
11. Weiner J, Parente S, Garnick D, Fowles J, Lawthers A, Palmer H: Variation in office based quality: a claims-based profile of care provided to Medicare patients with diabetes. *J Am Med Assoc* 273:1503–1508, 1995
12. Powe N, Weiner B, Starfield B, Stuart M, Baker A, Steinwachs D: Systemwide performance in a Medicaid program: profiling the care of patients with chronic illness. *Med Care* 34:798–810, 1996
13. Hofer TP: The unreliability of individual physician "report cards" for assessing the costs and quality of care of a chronic disease. *JAMA* 281:2098, 1999
14. Dudley RA: Pay for performance research: how to learn what clinicians and policy makers need to know. *J Am Med Assoc* 294:1821–1823, 2005
15. Rosenthal MB, Dudley RA: Pay for performance: will the latest payment trend improve care? *JAMA* 297:740–744, 2007
16. Krein S, Hofer T, Kerr E, Hayward R: Whom should we profile? Examining diabetes care practice variation among primary care providers, provider groups, and health care facilities. *Health Serv Res* 37:1159–1180, 2000
17. Miller DR, Safford MM, Pogach LM: Who has diabetes? Best estimates of diabetes prevalence in the Department of Veterans Affairs based on computerized patient data. *Diabetes Care* 27(Suppl. 2):B10–B21, 2004
18. American Diabetes Association: Conditions and treatment: A1C test. Available from <http://www.diabetes.org/type-1-diabetes/a1c-test.jsp>. Accessed 21 April 2007
19. Kreft I, DeLeeuw J. *Introducing Multilevel Modeling*. London, Sage Publications, 1998, p. 9
20. Raudenbush S, Bryk A, Cheong Y, Congdon R. *HLM6: Hierarchical Linear and Nonlinear Modeling*. Lincolnwood, IL, Scientific Software International, 2004
21. Bryk A, Raudenbush S. *Hierarchical Linear Models: Applications Analysis Methods*. New York, Sage Publications, 1992, p. 19
22. Snijders T, Bosker R: Modeling variance in two-level models. *Sociol Methods Res* 22:342–363, 1994
23. Tu YK, Githorpe MS: Revisiting the relation between change and initial value: a review and evaluation. *Stat Med* 26:443–457, 2007
24. Norris SL, Engelgau MM, Narayan KM: Effectiveness of self-management training in type 2 diabetes: a systematic review of randomized controlled trials. *Diabetes Care* 24:561–587, 2001
25. Norris SL, Lau J, Smith SJ, Schmid CH, Engelgau MM: Self-management education for adults with type 2 diabetes: a meta-analysis of the effect on glycemic control. *Diabetes Care* 25:1159–1171, 2002
26. Ismail K, Winkley K, Rabe-Hesketh S: Systematic review and meta-analysis of randomized controlled trials of psychological interventions to improve glycemic control in patients with type 2 diabetes. *Lancet* 363:1589–1597, 2004
27. Anderson RM, Funnell MM, Butler PM, Arnold MS, Fitzgerald JT, Feste CC: Patient empowerment: results of a randomized controlled trial. *Diabetes Care* 18:943–949, 1995
28. Martin L, Jahng K, Golin C, DiMatteo R: Physician facilitation of patient involvement in care: correspondence between patient and observer reports. *Behav Med* 28:159–164, 2003