



# Electronic Documentation of Lifestyle Counseling and Glycemic Control in Patients With Diabetes

*Diabetes Care* 2015;38:1326–1332 | DOI: 10.2337/dc14-2016

Naoshi Hosomura,<sup>1</sup> Saveli I. Goldberg,<sup>2</sup>  
 Maria Shubina,<sup>1</sup> Mary Zhang,<sup>1</sup> and  
 Alexander Turchin<sup>1,3,4</sup>

## OBJECTIVE

To establish quantitative characteristics of lifestyle counseling documentation associated with improved glycemic control in patients with diabetes.

## RESEARCH DESIGN AND METHODS

We retrospectively studied 10,870 hyperglycemic ( $\text{HbA}_{1c} \geq 7.0\%$  [53 mmol/mol]) adults with diabetes followed at primary care practices affiliated with two academic hospitals between 2000 and 2010. Documentation intensity was represented by the mean number of characters per note documenting lifestyle counseling. Heterogeneity was calculated as the normalized Levenshtein distance between lifestyle counseling sentences between consecutive notes. Cox proportional hazards model was constructed to assess association of heterogeneity and intensity of lifestyle counseling documentation to time to  $\text{HbA}_{1c} < 7.0\%$  (53 mmol/mol) while adjusting for demographics, initial  $\text{HbA}_{1c}$  level, insulin therapy, medication intensification, and frequency of lifestyle counseling.

## RESULTS

Comparing patients in the highest versus lowest tertile by documentation heterogeneity and documentation intensity, median time to  $\text{HbA}_{1c} < 7.0\%$  (53 mmol/mol) was 26 vs. 39 months and 24 vs. 39 months, respectively ( $P < 0.001$  for all). In multivariable analysis, an increase of documentation heterogeneity by 0.15 units and an increase of documentation intensity by 45 characters/note was associated with hazard ratios of 1.08 (95% CI 1.04–1.12;  $P < 0.001$ ) and 1.27 (95% CI 1.23–1.31;  $P < 0.001$ ) for time to  $\text{HbA}_{1c}$  target, respectively.

## CONCLUSIONS

Higher heterogeneity and intensity of lifestyle counseling documentation in provider notes were associated with better glycemic control. Further studies involving direct observation of patient care are needed to establish the nature of the relationship between documentation characteristics and patient outcomes.

The prevalence of diabetes is rapidly increasing in the U.S. and worldwide (1–3). Chronic hyperglycemia is associated with increased risk of microvascular and macrovascular complications, and maintaining glycemic levels as close to nondiabetic range as possible can substantially reduce the risk (4–6). Nevertheless, a large number of patients with diabetes do not reach target glycemic levels recommended by treatment guidelines (7–9).

<sup>1</sup>Division of Endocrinology, Brigham and Women's Hospital, Boston, MA

<sup>2</sup>Department of Radiation Oncology, Massachusetts General Hospital, Boston, MA

<sup>3</sup>Harvard Medical School, Boston, MA

<sup>4</sup>Harvard Clinical Research Institute, Boston, MA

Corresponding author: Alexander Turchin, aturchin@partners.org.

Received 21 August 2014 and accepted 28 March 2015.

This article contains Supplementary Data online at <http://care.diabetesjournals.org/lookup/suppl/doi:10.2337/dc14-2016/-/DC1>.

© 2015 by the American Diabetes Association. Readers may use this article as long as the work is properly cited, the use is educational and not for profit, and the work is not altered.

Randomized controlled trials have demonstrated beneficial effects of lifestyle interventions on glycemic control (10–12). Recent studies provide strong evidence that lifestyle counseling in routine care settings is associated with improved glycemic control, confirming the findings of clinical trials (13). However, evidence also suggests that lifestyle counseling practices in routine care settings are often less than optimal (14–16).

Federal incentives for meaningful use of certified electronic medical record (EMR) technology helped introduce a wide array of electronic measures of quality of medical care (17–19). Most of the proposed electronic measures (eMeasures) of lifestyle counseling provided to patients, however, focus on binary measures of the presence or absence of counseling during encounters rather than measures of counseling quality or effectiveness (20–22). This is in part because documentation characteristics of effective lifestyle counseling are not known.

The aim of this study was to establish quantitative characteristics of documentation of lifestyle counseling that are associated with improved glycemic control in patients with diabetes, using a previously validated natural language processing system that enables abstraction of lifestyle counseling documentation from narrative electronic provider notes (23). Previously published studies indicated that, unlike original records, copy-pasted documentation of lifestyle counseling was not associated with faster achievement of blood glucose control in patients with diabetes (23). It is also known that more intensive lifestyle counseling where patients are given detailed instructions on specific diet and exercise changes is more effective than less intensive methods (24–26). Based upon these findings, we hypothesized that more heterogeneous documentation of lifestyle counseling (where copy-paste documentation represents the lowest form of heterogeneity) and more intensive documentation of lifestyle counseling (which may be a reflection of more intensive counseling) are both associated with better patient outcomes. In order to test this hypothesis, we developed quantitative measures of heterogeneity and intensity of lifestyle counseling documentation in narrative electronic provider notes.

## RESEARCH DESIGN AND METHODS

### Design

We conducted a retrospective cohort study to determine whether quantitative characteristics of narrative electronic documentation of lifestyle counseling are associated with blood glucose control in patients with diabetes. We evaluated the relationship between two quantitative measures of documentation characteristics, documentation heterogeneity and documentation intensity, and time to HbA<sub>1c</sub> control.

### Study Cohort

Adult patients with diabetes followed by primary care physicians (PCPs) affiliated with Brigham and Women's Hospital (BWH) and Massachusetts General Hospital (MGH) between 1 January 2000 and 1 January 2010 were studied. Patients were included in the analysis if they were at least 18 years old, had a documented diagnosis of diabetes ("diabetes" on the EMR problem list) or HbA<sub>1c</sub> of at least 7.0% (53 mmol/mol), and had been followed for at least 2 years during the 10-year study period. Patients without at least one annual encounter with a PCP affiliated with BWH or MGH were excluded to eliminate those who were not actively treated in these practices. Patients with missing zip codes were excluded to enable adjustment for median household income by zip code. Demographic information, weight, height, and medication and laboratory data were obtained from the EMR at Partners HealthCare, an integrated health care delivery network in eastern Massachusetts that includes BWH and MGH. During the study period, none of the study practices had a program that encouraged a particular type of lifestyle counseling or monitored lifestyle counseling delivered by providers. The EMR system had a number of decision support features but did not include any decision support for lifestyle counseling.

This study was approved by the Partners HealthCare System institutional review board, and the requirement for written informed consent was waived.

### Study Measurements

A single "hyperglycemic" period served as the unit of analysis. A hyperglycemic period started at the first available HbA<sub>1c</sub> measurement of 7.0% (53 mmol/mol) or greater and ended at

the first HbA<sub>1c</sub> measurement <7.0% (53 mmol/mol) or at the end of the study period if HbA<sub>1c</sub> never reached treatment target (27). Time to HbA<sub>1c</sub> control was defined as the length of the hyperglycemic period. Each patient may have contributed multiple hyperglycemic periods if HbA<sub>1c</sub> measurements fluctuated above and below the target level during the study period.

Hyperglycemic periods without any medication information available in the EMR were excluded to enable inclusion of insulin treatment as a confounder variable in the analysis. The medication information was based upon EMR prescription records. We excluded patients with two or more visits to an endocrinologist in order to focus our analyses on the primary care setting, where the large majority of patients with diabetes are treated. Periods in which the rate of change of HbA<sub>1c</sub> was >3 SD from the mean were excluded to eliminate likely measurement errors. Transient elevations of HbA<sub>1c</sub> were defined as periods that contained only a single elevated measurement followed by a fall below the target level at the next measurement with no interceding antihyperglycemic medication intensification, and these periods were excluded from the analysis. Medication intensification rate was defined as the mean monthly number of episodes of antihyperglycemic medication intensification, indicated by initiation of a new medication or an increase in the dose of an existing medication. Medication intensification was abstracted from a combination of structured medication records and computational analysis of narrative provider notes as previously described (28). Frequency of lifestyle counseling was calculated as the mean monthly number of encounters during the hyperglycemic period in which a PCP documented diet, exercise, or weight loss counseling.

The large majority, but not all, of the notes were generated by PCPs, some of whom were staff physicians and others were trainees. While we were unable to distinguish between notes by PCPs versus notes by nonphysician clinicians including dietitians, there were very few advanced practice providers in BWH/MGH at the time when the data used in the study were collected.

Natural language processing is a technology that enables rapid extraction of

specific clinical information through computational analysis of clinical narratives. In the current study, sentences dedicated to documenting lifestyle counseling were automatically extracted from narrative EMR notes using previously validated natural language processing software. The software identified three categories of lifestyle counseling: diet counseling, exercise counseling, and weight loss counseling, with a sensitivity and specificity that ranged between 91 and 97% and 88 and 94%, respectively (23).

Documentation heterogeneity was calculated for two consecutive notes with documented lifestyle counseling as the Levenshtein distance between the relevant sentences, normalized by the length of the longer sentence. The Levenshtein distance between two strings is a metric that represents how similar the two strings are and is defined as the minimum number of editing operations needed to transform one string into the other (29). Examples of Levenshtein distance calculation are illustrated in Supplementary Figs. 1 and 2. In order to quantify the heterogeneity of lifestyle counseling documentation, we applied this concept to the note level. (A detailed description of the algorithm can be found in the Supplementary Data.) Sample data and calculation results of documentation heterogeneity can be found in Supplementary Fig. 3. We empirically found that longer sentences frequently were composed of multiple sentences without separating punctuation and therefore did not provide a good representation of lifestyle counseling documentation heterogeneity. We therefore excluded sentences longer than 100 characters in calculating documentation heterogeneity. Since it has been shown that copied documentation of lifestyle counseling is not associated with improvement in glucose control (23), we excluded hyperglycemic periods in which the lifestyle counseling documentation consisted entirely of copied/duplicate sentences. If the normalized Levenshtein distance is zero, this means that the relevant sentences being compared are all identical, indicating that the lifestyle counseling documentation consisted entirely of copied sentences throughout the hyperglycemic period. We therefore excluded hyperglycemic

periods whose normalized Levenshtein distance equaled to zero.

By measuring “heterogeneity” of documentation, we aimed at quantifying the textual diversity of the documentation between consecutive lifestyle counseling notes. When lifestyle counseling episodes are documented using similar or identical sentences between consecutive counseling episodes, such hyperglycemic periods would be characterized to have low documentation heterogeneity. Copy-pasted documentation of lifestyle counseling would represent an extreme form of low heterogeneity. On the other hand, when the sentences used to document counseling episodes are more dissimilar to one another between consecutive notes, such hyperglycemic periods would be characterized to have high documentation heterogeneity.

Documentation intensity was calculated as the mean number of characters per note dedicated to documenting lifestyle counseling. Although it was possible to calculate the documentation intensity for the hyperglycemic periods for which documentation heterogeneity was incalculable (i.e., periods with only one note with documented lifestyle counseling per counseling category), we used the same set of hyperglycemic periods for both predictor variables in our analyses in order to assess relative contributions of documentation heterogeneity and documentation intensity.

#### Statistical Analysis

Summary statistics were conducted by using frequencies and proportions for categorical data and using means, SDs, and medians for continuous variables. In univariate analyses, hyperglycemic periods were grouped into tertiles by each predictor variable for ease of interpretation. The log-rank test was used to compare Kaplan-Meier cumulative incidence curves for time to HbA<sub>1c</sub> target between the tertile groups. Marginal Cox proportional hazards regression model for clustered data (29) was used to estimate the association between time to HbA<sub>1c</sub> control and documentation characteristics while accounting for clustering within individual patients and adjusting for demographic confounders (age, sex, race/ethnicity, primary language, health insurance, and median income by zip code), Charlson comorbidity index, presence of obesity (defined as documented BMI

≥30 kg/m<sup>2</sup>) during the period, frequency of encounters with documented lifestyle counseling, frequency of medication intensification, frequency of HbA<sub>1c</sub> measurement, and the HbA<sub>1c</sub> level at the start of the hyperglycemic period. When adjusting for clustering within individual providers, the patient’s PCP was defined as the physician in a primary care practice who contributed the highest number of notes used for calculation of documentation characteristics during the hyperglycemic period.

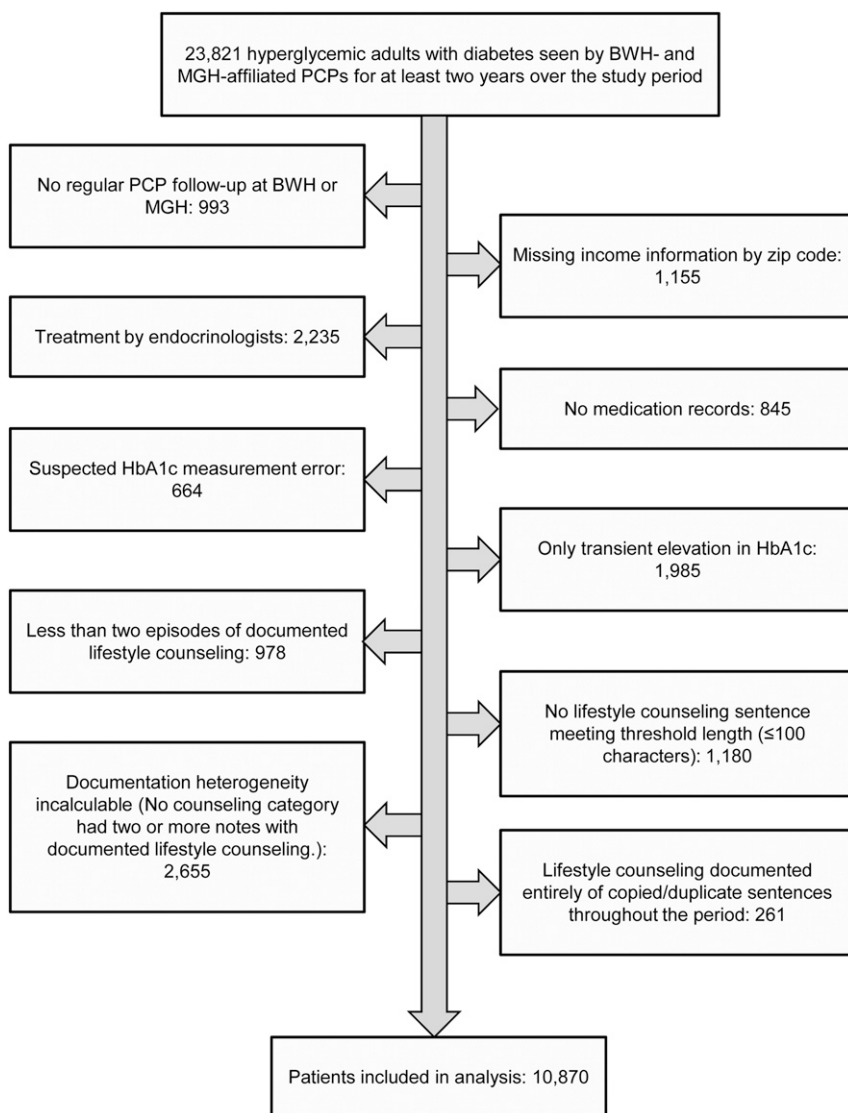
*P* values were adjusted for multiple hypothesis testing using the Simes-Hochberg method (30,31). All the analyses were performed using SAS, version 9.3 (SAS Institute, Cary, NC).

#### RESULTS

We identified 23,821 adults with diabetes followed by BWH- or MGH-affiliated PCPs for at least 2 years over the study period (Fig. 1). After exclusion of patients without regular PCP follow-up, with missing income information by zip code, regularly treated by endocrinologists, with no medication records, with suspected HbA<sub>1c</sub> measurement errors, with only transient elevations in HbA<sub>1c</sub>, and patients for whom lifestyle counseling documentation metrics could not be calculated, the remaining 10,870 patients, who contributed to a total of 13,594 hyperglycemic periods, were included in the study.

Study patients did not have their HbA<sub>1c</sub> under control over a mean of 50% of total follow-up time. Their mean initial HbA<sub>1c</sub> level at the beginning of hyperglycemic period was 8.3% (Table 1). The included hyperglycemic periods had a median length of 32 months. Lifestyle counseling was documented at a mean rate of once in every 3.3 months. The mean rate of HbA<sub>1c</sub> testing was once in 3.6 months. In ~40% of the hyperglycemic periods, patients never achieved target HbA<sub>1c</sub> level during the study period (Table 2).

A total of 183,611 sentences describing lifestyle counseling from 92,671 provider notes were analyzed to calculate the heterogeneity and intensity of lifestyle counseling documentation. The mean documentation heterogeneity, represented by the normalized Levenshtein distance, was 0.69. The mean documentation intensity was 91 characters per note.



**Figure 1**—Exclusion criteria and numbers of patients excluded.

Time to achievement of target HbA<sub>1c</sub> level decreased progressively both with increasing heterogeneity and with increasing intensity of lifestyle counseling documentation (Fig. 2). Compared with hyperglycemic periods in the highest tertile by documentation heterogeneity and intensity, median time to HbA<sub>1c</sub> control in the lowest tertile by documentation heterogeneity and intensity was 26 vs. 39 months and 24 vs. 39 months, respectively ( $P < 0.001$  for all).

In multivariable Cox proportional hazards models adjusted for the patients' demographic characteristics, presence of obesity during the period, Charlson comorbidity index, treatment with insulin, HbA<sub>1c</sub> level at the start of the period, frequency of HbA<sub>1c</sub> measurement, lifestyle counseling frequency, frequency of

medication intensification, and clustering within individual patients, 1-SD (0.15 units) increase in documentation heterogeneity, and 1-SD (45 characters per note) increase in documentation intensity were associated with hazard ratios of 1.08 (95% CI 1.04–1.12;  $P < 0.001$ ) and 1.27 (95% CI 1.23–1.31;  $P < 0.001$ ) for time to HbA<sub>1c</sub> control, respectively (Table 3). The notes evaluated in the analyses were generated by 1,155 providers. The relationship of documentation heterogeneity and intensity with time to HbA<sub>1c</sub> control did not change substantially when the multivariable analysis was adjusted for clustering within individual providers (Supplementary Table 1).

We also conducted a sensitivity analysis that included previously excluded

hyperglycemic periods (no income information, no medication information, treatment by endocrinologists, and only transient elevations of HbA<sub>1c</sub>), where provider age and sex were included as additional covariates. The data set included 19,562 hyperglycemic periods contributed by 14,863 unique patients. The results showed that neither provider characteristic had a significant relationship with time to HbA<sub>1c</sub> target ( $P = 0.08$  and  $P = 0.51$  for provider age and sex, respectively), and the relationship of documentation heterogeneity and intensity with time to HbA<sub>1c</sub> control did not substantially change (Supplementary Table 1). Among the patients included in the analyses, 9,766 of 10,870 (90%) had diagnosed diabetes. We have conducted a sensitivity analysis limited to the patients with diagnosed diabetes, which did not substantially change the results (Supplementary Table 1). To evaluate whether there was variability in results over time, we performed a sensitivity analysis for two separate data sets obtained by splitting the original data set into halves by a cut-off date (hyperglycemic period start date of 31 December 2004), which did not substantially change the results (Supplementary Table 1). An additional multivariable analysis was conducted to evaluate the potential interactions between obesity versus documentation heterogeneity and documentation intensity. The results showed that obesity was not an effect modifier of either of the documentation characteristics in their association with time to HbA<sub>1c</sub> control ( $P = 0.96$  and  $P = 0.45$  for documentation heterogeneity and documentation intensity, respectively). Lastly, for evaluation of the possibility that the documentation characteristics reflect processes other than lifestyle counseling affecting time to HbA<sub>1c</sub> control, correlation analyses were conducted to assess the relationship between frequency of medication intensification and each predictor variable. The results showed only limited correlations between these variables (Pearson correlation coefficient 0.027 for documentation heterogeneity and 0.030 for documentation intensity).

**CONCLUSIONS**

In this large, long-term retrospective study, we identified novel quantitative

**Table 1—Descriptive statistics of study patients' demographic and clinical characteristics**

N	10,870
Age, mean (SD), years <sup>a</sup>	59.3 (13.5)
Women, n (%)	5,635 (51.8)
Race/ethnicity, n (%)	
White	6,447 (59.3)
Black	1,576 (14.5)
Hispanic	1,772 (16.3)
Asian	370 (3.4)
Other <sup>b</sup>	705 (6.5)
English as the primary language, n (%)	8,623 (79.3)
Health insurance, n (%)	
Private	4,336 (39.9)
Medicare	5,162 (47.5)
Medicaid	1,183 (10.9)
None/unknown	189 (1.7)
Median income by zip code, mean (SD), \$1000s	50.5 (19.5)
No. of hyperglycemic periods, mean (SD)	1.3 (0.6)
HbA <sub>1c</sub> , mean (SD), %; mmol/mol	8.3 (1.7); 67 (13.7)
Charlson comorbidity index, mean (SD)	6.4 (4.6)
Follow-up time, mean (SD), months	82.2 (30.6)
Total time above treatment target, mean (SD), months	41.5 (30.7)

<sup>a</sup>Age calculated at the start date of the first hyperglycemic period. <sup>b</sup>Includes unknown.

characteristics of electronic documentation of lifestyle counseling that are associated with improved glycemic control in patients with diabetes. To our knowledge, this is the first study to describe an association between metrics of documentation characteristics and patient outcomes through a large-scale quantitative analysis of narrative EMR data.

While the retrospective nature of this study does not allow us to establish a direct causal relationship, there are several plausible explanations for the associations between the metrics of

documentation characteristics and patients' glycemic control. For example, intensity of documentation of lifestyle counseling could reflect the extent of the counseling provided. Providers who spend more time counseling the patient and provide more detailed instructions may also document the counseling episode in greater detail, the extent of which could be represented by documentation intensity. Documentation heterogeneity, on the other hand, could reflect the degree of the variety of advice or the amount of new information content of counseling between

different provider-patient encounters. Trying a different diet or exercise approach after the previous one did not work, for example, may represent an effective lifestyle counseling strategy. The reverse dose-response relationship between documentation heterogeneity and time to HbA<sub>1c</sub> control is consistent with the previously published findings that copy-pasted documentation of lifestyle counseling was not associated with improvements in blood glucose control (23).

The metrics we studied could be utilized in several ways. They could be used, for example, to identify healthcare providers who may be less skilled in the complicated art of lifestyle counseling and could benefit from additional training. They could also be used to identify the most skilled providers who might be best able to deliver effective lifestyle counseling for particularly complicated patients. On the other hand, the metrics would likely be less useful for direct feedback to providers, in part because of the risk of "documenting to the metric."

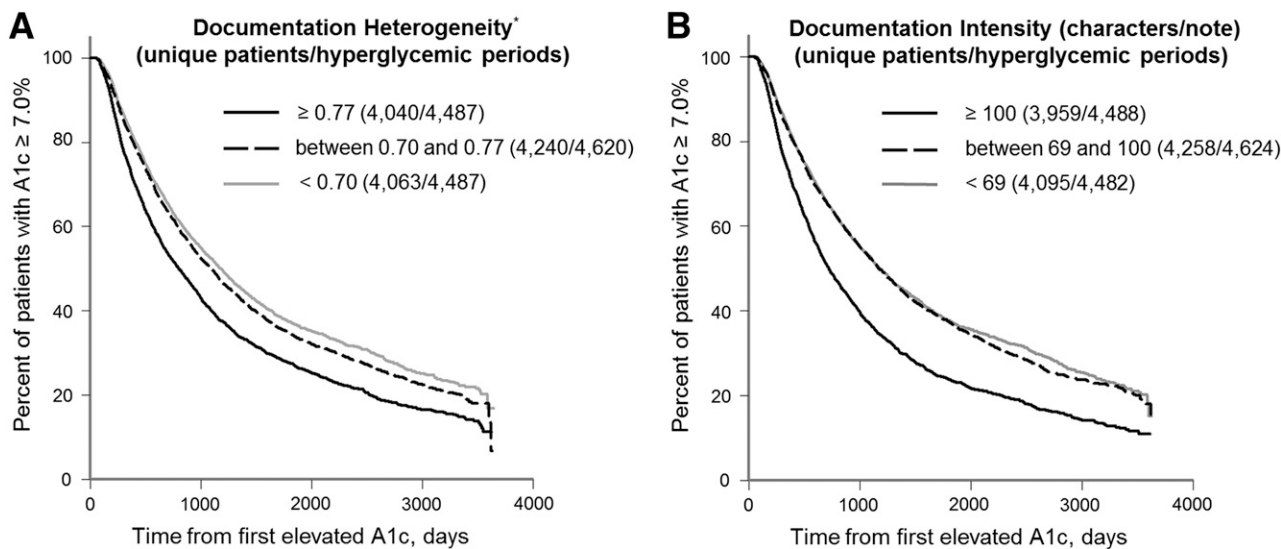
A key challenge in leveraging information from EMRs to its full capacity is the predominance of unstructured data in clinical documentation (32). This study showcases the power that EMRs combined with advanced technologies, such as natural language processing, bring to measurement of quality of care. Meaningful Use and other federal programs are rapidly expanding adoption of eMeasures—quality measures that are computed using EMRs data (33). Our study shows that it is possible to push the horizons of e-quality measurement even further, beyond structured clinical and administrative data, to take advantage of the wealth of information in the narrative electronic documents. The tools and approaches we developed can potentially be used in other health care settings to identify documentation characteristics that may reflect the underlying quality of care.

The results of this study need to be interpreted in the light of its limitations. Approximately half of the patients who met the initial inclusion criteria had to be excluded from the primary analyses. This could potentially lead to a bias if the relationship between documentation characteristics and time to HbA<sub>1c</sub> control was markedly different in the patient population excluded from the

**Table 2—Descriptive statistics of hyperglycemic period characteristics**

N	13,594
Period length, mean (SD), months	33.2 (28.1)
Initial HbA <sub>1c</sub> , mean (SD), %; mmol/mol	8.3 (1.6); 67 (12.9)
Maximum HbA <sub>1c</sub> , mean (SD), %; mmol/mol	9.4 (2.0); 79 (16.8)
Periods where treatment target was reached, n (%)	8,300 (61.1)
Rate of HbA <sub>1c</sub> measurement per month <sup>-1</sup> , mean (SD)	0.28 (0.16)
Rate of lifestyle counseling per month <sup>-1</sup> , mean (SD)	0.30 (0.24)
Rate of medication intensification per month <sup>-1</sup> , mean (SD)	0.10 (0.13)
Periods with patients receiving insulin, n (%)	4,118 (30.3)
Periods with patients who are obese (BMI ≥30 kg/m <sup>2</sup> ), n (%)	7,549 (55.5)
Documentation heterogeneity, mean (SD) <sup>a</sup>	0.69 (0.15)
Documentation intensity, mean (SD), characters/note	91 (45)

<sup>a</sup>Documentation heterogeneity represented by the normalized Levenshtein distance.



**Figure 2**—Kaplan-Meier curves for documentation characteristics and time to HbA<sub>1c</sub> control. *A*: Documentation heterogeneity and time to HbA<sub>1c</sub> target. (\*Documentation heterogeneity represented by the normalized Levenshtein distance.) *B*: Documentation intensity and time to HbA<sub>1c</sub> target (*P* < 0.001 by log-rank test for all).

analyses. Since calculation of documentation heterogeneity by definition required two or more notes with documented lifestyle counseling, documentation heterogeneity would not be applicable in case a patient receives only one episode of lifestyle counseling. We excluded the hyperglycemic periods

in which lifestyle counseling documentation consisted entirely of copied/duplicate sentences, as inclusion of these hyperglycemic periods would likely confound the study results. We were unable to distinguish between patients with type 1 versus type 2 diabetes. The majority of patients in the study

population had type 2 diabetes, so our findings may not be applicable to patients with type 1 diabetes. Another limitation is that we were unable to distinguish between notes by PCPs and notes by nonphysician clinicians. There were very few nonphysician clinicians at BWH/MGH at the time when the data used in the study were collected, so there was little team-based care. Further studies are needed to evaluate whether our findings would be applicable when a larger portion of notes are generated by nonphysician clinicians. The accuracy of the natural language processing software was not perfect, although the sensitivity and specificity of the software were very high. This study was conducted at two academic hospitals in eastern Massachusetts, and this could limit the generalizability of the study results to other practice settings. Lastly, the retrospective nature of this study does not allow us to make causal inferences in the associations that we have found.

In conclusion, this large, long-term retrospective study identified novel quantitative characteristics of electronic documentation of lifestyle counseling that are associated with improved glycemic control in patients with diabetes. Both higher heterogeneity and higher intensity of lifestyle counseling documentation in narrative provider notes were associated with faster achievement of target HbA<sub>1c</sub> levels. Further studies involving direct observation of patient care

**Table 3**—Hazard ratios for associations between documentation characteristics and time to HbA<sub>1c</sub> control<sup>a</sup>

Variable	Hazard ratio (95% CI)	<i>P</i> <sup>b</sup>
Female sex	0.92 (0.87–0.98)	0.005
White race (vs. all nonwhites)	0.97 (0.92–1.04)	0.40
English speaker	1.02 (0.95–1.09)	0.57
Income, per \$1,000 increase	0.998 (0.997–1.00)	0.03
Government insurance	0.98 (0.92–1.04)	0.41
Insulin treatment	0.38 (0.35–0.41)	<0.001
Obesity during period	0.99 (0.92–1.05)	0.69
Charlson comorbidity index	1.01 (1.00–1.02)	<0.001
Rate of HbA <sub>1c</sub> testing, per 3 months <sup>-1</sup>	3.58 (2.94–4.35)	<0.001
Rate of medication intensification, per month <sup>-1</sup>	3.39 (2.19–5.26)	<0.001
Rate of lifestyle counseling, per month <sup>-1</sup>	7.70 (5.93–10.00)	<0.001
HbA <sub>1c</sub> level at start of period, per 1% increase	0.85 (0.83–0.86)	<0.001
Age, per 10-year increase	1.07 (1.04–1.10)	<0.001
Documentation heterogeneity, per 1-SD increase (0.15 units) <sup>c</sup>	1.06 (1.04–1.12)	<0.001
Documentation intensity, per 1-SD increase (45 characters/note)	1.27 (1.23–1.31)	<0.001

Data are hazard ratios as estimated by multivariable Cox regression model adjusted for patient and treatment characteristics. <sup>a</sup>Values are reported as hazard ratios for reaching an HbA<sub>1c</sub> value <7.0%. <sup>b</sup>*P* values <0.001 are significant when adjusted for multiple hypothesis testing using the Simes-Hochberg method. <sup>c</sup>Documentation heterogeneity measured as the normalized Levenshtein distance.

are needed to establish the nature of the relationship between documentation characteristics and patient outcomes.

**Acknowledgments.** The authors thank Dr. Anna Rumshisky, PhD, at the Department of Computer Science, University of Massachusetts Lowell. Dr. Rumshisky contributed to study concept and design, and important intellectual content of the study and did not receive any compensation for her contributions.

**Funding.** This study was supported in part by grant 1R18HS017030 from the Agency for Healthcare Research and Quality (to S.I.G., M.S., and A.T.).

The funding sources had no direct impact on the design and conduct of the study; the collection, management, analysis, and interpretation of the data; or the preparation, review, or approval of the manuscript.

**Duality of Interest.** No potential conflicts of interest relevant to this article were reported.

**Author Contributions.** N.H., S.I.G., M.S., M.Z., and A.T. analyzed data. N.H. drafted the manuscript. N.H., S.I.G., M.S., M.Z., and A.T. critically revised the manuscript for important intellectual content. N.H., S.I.G., M.S., and A.T. performed statistical analysis. S.I.G., M.S., and A.T. developed the study concept and design. M.Z. provided administrative, technical, and material support. A.T. acquired data, obtained funding, and supervised the study. A.T. is the guarantor of this work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

**Prior Presentation.** Parts of this study were presented in abstract form at the 75th Scientific Sessions of the American Diabetes Association, Boston, MA, 5–9 June 2015.

## References

- U.S. Department of Health and Human Services. *National Diabetes Statistics Report: Estimates of Diabetes and Its Burden in the United States, 2014*. Atlanta, GA, Centers for Disease Control and Prevention, 2014
- Guariguata L, Whiting DR, Hambleton I, Beagley J, Linnenkamp U, Shaw JE. Global estimates of diabetes prevalence for 2013 and projections for 2035. *Diabetes Res Clin Pract* 2014; 103:137–149
- Danaei G, Finucane MM, Lu Y, et al.; Global Burden of Metabolic Risk Factors of Chronic Diseases Collaborating Group (Blood Glucose). National, regional, and global trends in fasting plasma glucose and diabetes prevalence since 1980: systematic analysis of health examination surveys and epidemiological studies with 370 country-years and 2·7 million participants. *Lancet* 2011;378:31–40
- Coca SG, Ismail-Beigi F, Haq N, Krumholz HM, Parikh CR. Role of intensive glucose control in development of renal end points in type 2 diabetes mellitus: systematic review and meta-analysis intensive glucose control in type 2 diabetes. *Arch Intern Med* 2012;172:761–769
- Selvin E, Coresh J, Golden SH, Brancati FL, Folsom AR, Steffes MW. Glycemic control and coronary heart disease risk in persons with and without diabetes: the atherosclerosis risk in communities study. *Arch Intern Med* 2005; 165:1910–1916
- Patel A, MacMahon S, Chalmers J, et al.; ADVANCE Collaborative Group. Intensive blood glucose control and vascular outcomes in patients with type 2 diabetes. *N Engl J Med* 2008;358:2560–2572
- Ali MK, Bullard KM, Saaddine JB, Cowie CC, Imperatore G, Gregg EW. Achievement of goals in U.S. diabetes care, 1999–2010. *N Engl J Med* 2013;368:1613–1624
- Stark Casagrande S, Fradkin JE, Saydah SH, Rust KF, Cowie CC. The prevalence of meeting A1C, blood pressure, and LDL goals among people with diabetes, 1988–2010. *Diabetes Care* 2013;36:2271–2279
- Selvin E, Parrinello CM, Sacks DB, Coresh J. Trends in prevalence and control of diabetes in the United States, 1988–1994 and 1999–2010. *Ann Intern Med* 2014;160:517–525
- Knowler WC, Barrett-Connor E, Fowler SE, et al.; Diabetes Prevention Program Research Group. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346:393–403
- Nield L, Moore HJ, Hooper L, et al. Dietary advice for treatment of type 2 diabetes mellitus in adults. *Cochrane Database Syst Rev* 2007; CD004097
- Thomas DE, Elliott EJ, Naughton GA. Exercise for type 2 diabetes mellitus. *Cochrane Database Syst Rev* 2006; CD002968
- Morrison F, Shubina M, Turchin A. Lifestyle counseling in routine care and long-term glucose, blood pressure, and cholesterol control in patients with diabetes. *Diabetes Care* 2012; 35:334–341
- Galuska DA, Will JC, Serdula MK, Ford ES. Are health care professionals advising obese patients to lose weight? *JAMA* 1999;282: 1576–1578
- Egede LE, Zheng D. Modifiable cardiovascular risk factors in adults with diabetes: prevalence and missed opportunities for physician counseling. *Arch Intern Med* 2002;162: 427–433
- Ma J, Urizar GG Jr, Alehegn T, Stafford RS. Diet and physical activity counseling during ambulatory care visits in the United States. *Prev Med* 2004;39:815–822
- Marcotte L, Seidman J, Trudel K, et al. Achieving meaningful use of health information technology: a guide for physicians to the EHR incentive programs. *Arch Intern Med* 2012; 172:731–736
- Weiner JP, Fowles JB, Chan KS. New paradigms for measuring clinical performance using electronic health records. *Int J Qual Health Care* 2012;24:200–205
- Garrido T, Kumar S, Lekas J, et al. e-Measures: insight into the challenges and opportunities of automating publicly reported quality measures. *J Am Med Assoc* 2014;21:181–184
- Shaikh U, Nettiksimmons J, Bell RA, Tancredi D, Romano PS. Accuracy of parental report and electronic health record documentation as measures of diet and physical activity counseling. *Acad Pediatr* 2012;12:81–87
- Hazlehurst BL, Lawrence JM, Donahoo WT, et al. Automating assessment of lifestyle counseling in electronic health records. *Am J Prev Med* 2014;46:457–464
- Zimmermann LJ, Thompson JA, Persell SD. Electronic health record identification of prediabetes and an assessment of unmet counseling needs. *J Eval Clin Pract* 2012;18:861–865
- Turchin A, Goldberg SI, Breydo E, Shubina M, Einbinder JS. Copy/paste documentation of lifestyle counseling and glycemic control in patients with diabetes: true to form? *Arch Intern Med* 2011;171:1393–1394
- Greaves CJ, Sheppard KE, Abraham C, et al.; IMAGE Study Group. Systematic review of reviews of intervention components associated with increased effectiveness in dietary and physical activity interventions. *BMC Public Health* 2011;11:119
- Costa B, Barrio F, Cabré J-J, et al.; DE-PLAN-CAT Research Group. Delaying progression to type 2 diabetes among high-risk Spanish individuals is feasible in real-life primary healthcare settings using intensive lifestyle intervention. *Diabetologia* 2012; 55:1319–1328
- Gregg EW, Chen H, Wagenknecht LE, et al.; Look AHEAD Research Group. Association of an intensive lifestyle intervention with remission of type 2 diabetes. *JAMA* 2012;308:2489–2496
- Morrison F, Shubina M, Turchin A. Encounter frequency and serum glucose level, blood pressure, and cholesterol level control in patients with diabetes mellitus. *Arch Intern Med* 2011;171:1542–1550
- Turchin A, Shubina M, Breydo E, Pendergrass ML, Einbinder JS. Comparison of information content of structured and narrative text data sources on the example of medication intensification. *J Am Med Inform Assoc* 2009;16: 362–370
- Lin DY. Cox regression analysis of multivariate failure time data: the marginal approach. *Stat Med* 1994;13:2233–2247
- Simes RJ. An improved Bonferroni procedure for multiple tests of significance. *Biometrika* 1986;73:751–754
- Hochberg Y. A sharper Bonferroni procedure for multiple tests of significance. *Biometrika* 1988;75:800–802
- Agency for Healthcare Research and Quality. Challenges of measuring care coordination using electronic data and recommendations to address those challenges: prospects for care coordination measurement using electronic data sources [Internet], 2013. Available from <http://www.ahrq.gov/research/findings/final-reports/prospectscare/prospectscare1.html>. Accessed 19 July 2014
- National Quality Forum. Measuring performance: electronic quality measures (eMeasures) [Internet], 2012. Washington, DC, National Quality Forum. Available from [http://www.qualityforum.org/projects/e-g/emeasures/electronic\\_quality\\_measures.aspx](http://www.qualityforum.org/projects/e-g/emeasures/electronic_quality_measures.aspx). Accessed 19 July 2014