Electronic Documentation as a Reflection of Quality of Care

A Thesis Presented by

Naoshi Hosomura

to

The Faculty of Medicine

In partial fulfillment of the requirements

for the degree of

Doctor of Medical Sciences

Research Mentor: Alexander Turchin, MD, MS

Associate Professor of Medicine, Harvard Medical School

Division of Endocrinology, Brigham and Women’s Hospital

Director of Clinical Informatics, Harvard Clinical Research Institute

Harvard School of Dental Medicine

Boston, Massachusetts

May 2014
Acknowledgements

I would like to express the deepest appreciation to my research mentor, Dr. Alexander Turchin, who has provided me with all the necessary resources to make this dissertation possible. I am extremely grateful and deeply indebted to him for his continuous support, sincere guidance, patience, and understanding throughout my thesis research. Without his thoughtful encouragement and careful supervision, this thesis would never have taken shape.

I am grateful to Dr. Saveli I Goldberg, who was instrumental in conceiving this study and provided many insightful suggestions. I would like to thank Dr. Maria Shubina, who also provided invaluable advice and suggestions. I would like to thank Mary Zhang who kindly shared analytical code that served as a starting point for this project. I also would like to thank Dr. Anna Rumshisky for her essential contributions in the early design stages of this study, as well as her helpful feedback in the late stages of this work.

I would like to thank the members of my Thesis Advisory Committee, Dr. Steven Simon, Dr. Thomas Sequist, and Dr. Kenneth Mandl. Their thoughtful questions and comments were valued greatly.

I also would like to thank the members of my Thesis Defense Committee, Dr. Dennis Ross-Degnan, Dr. Lipika Samal, and Dr. Steven Simon who also served on the Advisory Committee, for their careful review of this manuscript and helpful feedback that has been incorporated into this thesis.
## Table of Contents

Title Page ......................................................................................... 1
Signature Page ................................................................................ 2
Copyright Statement ........................................................................ 3
Acknowledgements .......................................................................... 4
Table of Contents ........................................................................... 5
Abstract ......................................................................................... 8
Hypothesis and Specific Aims ........................................................... 11

### Research Strategy

- Significance .................................................................................. 13
- Innovation ...................................................................................... 17
- Approach ....................................................................................... 20
  - Design .......................................................................................... 20
  - Data ............................................................................................. 20
  - Study Cohort ............................................................................... 21
  - Study Measurements ................................................................. 22
- Statistical Analysis ........................................................................ 32

### Results

- ........................................................................................................ 34

### Discussion

- ........................................................................................................ 46

### Conclusions

- ........................................................................................................ 50

### References

- ........................................................................................................ 51
List of Figures and Tables

Figures

Figure 1. Illustration of a hyperglycemic period ........................................ 23
Figure 2. Algorithm for calculation of documentation heterogeneity ............ 26
Figure 3. Example of documentation heterogeneity calculation ................... 27
Figure 4. Algorithm for calculation of documentation intensity ................... 30
Figure 5. Example of documentation intensity calculation ........................ 31
Figure 6. Sample data and calculation results of documentation heterogeneity and documentation intensity ................................................................. 32
Figure 7. Schematic representation of the study protocol ............................. 34
Figure 8. Counts of hyperglycemic periods excluded from the analysis .......... 36
Figure 9. Documentation characteristics and time to A1c target .................. 39
Figure 10. Schematic representation of the hypothesized relationship between documentation characteristics and clinical outcomes in the Chronic Care Model (CCM) framework ................................................................. 48

Tables

Table 1. Patient characteristics ............................................................... 37
Table 2. Hyperglycemic period characteristics .......................................... 38
Table 3. Effects of patient, treatment, and documentation characteristics on time to A1c control accounting for clustering within individual patients ................................................................. 41
Table 4. *Effects of patient, treatment, and documentation characteristics on time to A1c control accounting for clustering within individual providers*. 42

Table 5. *Correlations between the confounder variables and documentation heterogeneity*. 43

Table 6. *Correlations between the confounder variables and documentation intensity*. 44

Table 7. *Effects of patient, provider, treatment, and documentation characteristics on time to A1c control accounting for clustering within individual patients*. 45
**ABSTRACT**

**Objectives**

While documentation is the most basic functionality of any Electronic Medical Record system, and majority of clinicians’ time is spent gathering and recording information, little empirical research has focused on evaluating how specific quantitative characteristics of electronic documentation may be associated with patient outcomes. We tested a hypothesis that quantifiable characteristics of documentation of lifestyle counseling in narrative provider notes are associated with blood glucose control in patients with diabetes.

**Methods**

We performed a retrospective study of adult patients with diabetes treated at primary care practices affiliated with 2 academic hospitals between 01/01/2000 and 01/01/2011. A single hyperglycemic period, defined as a period of continuously elevated blood glucose level (hemoglobin A1c ≥7.0%), served as the unit of analysis. Documentation of lifestyle counseling was extracted automatically from narrative provider notes using a previously validated natural language processing system. We quantified the characteristics of lifestyle counseling documentation by using the following two measures: 1) *documentation heterogeneity*, 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, and 2)
documentation intensity, represented by the mean number of characters per note
dedicated to documenting lifestyle counseling. The relationship between these
documentation characteristics and time to achievement of target A1c level (<7%) was
assessed.

Results

A total of 13,549 distinct hyperglycemic periods from 10,870 unique patients
were included in the analysis. 183,611 lifestyle counseling sentences from 92,671
provider notes were studied to calculate the heterogeneity and intensity of lifestyle
counseling documentation. Comparing hyperglycemic periods in the highest versus
lowest tertile by documentation heterogeneity and documentation intensity, median time
to A1c < 7.0% was 26 versus 39 months, and 24 versus 39 months, respectively (P <
0.0001 for all). In multivariable analysis adjusted for the patients’ demographic
characteristics, initial A1c level, frequency of A1c measurement, treatment with insulin,
frequency of medication intensification, frequency of encounters with documented
lifestyle counseling, Charlson comorbidity index, presence of obesity during the period,
and clustering within individual patients, an increase of documentation heterogeneity by
1 unit and an increase of documentation intensity by 100 characters/note was associated
with hazard ratios of 1.68 (95% CI 1.33 to 2.12; P < 0.0001) and 1.70 (95% CI 1.59 to
1.81; P < 0.0001) for time to achievement of A1c target, respectively.

Conclusions
This large long-term retrospective study found that both higher documentation heterogeneity and higher documentation intensity of lifestyle counseling are associated with faster achievement of A1c control. The findings suggest that these two quantitative characteristics of lifestyle counseling documentation may reflect the quality, quantity, and/or effectiveness of lifestyle counseling provided. Heterogeneity and intensity of documentation of lifestyle counseling in narrative provider notes may be used as indicators of quality of diabetes care.
HYPOTHESIS AND SPECIFIC AIMS

Hypothesis

Quantitative characteristics of documentation of lifestyle counseling in narrative provider notes reflect the quality, quantity, and/or effectiveness of the counseling provided and therefore are associated with differences in patient outcomes affected by the counseling.

Specific Aims

Specific Aim 1

To determine whether the heterogeneity of documentation of lifestyle counseling is associated with blood glucose control in patients with diabetes

To test our hypothesis that quantitative characteristics of narrative electronic documentation are associated with patient outcomes, we developed a measure termed documentation heterogeneity. 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. Our supposition was that documentation of counseling episodes in which providers discussed more diverse aspects of lifestyle intervention or provided greater amount of new information would be characterized by higher documentation heterogeneity. We assessed
the relationship between documentation heterogeneity and time to achievement of target A1c level (<7%).

**Specific Aim 2**

*To determine whether the intensity of documentation of lifestyle counseling is associated with blood glucose control in patients with diabetes mellitus*

We developed a measure termed *documentation intensity*, represented by the mean number of characters per note dedicated to documenting lifestyle counseling. We hypothesized that documentation of counseling episodes in which providers spent more time counseling the patient and/or provided more intensive, detailed lifestyle counseling would be characterized by higher documentation intensity. We assessed the relationship between documentation intensity and time to achievement of target A1c level.
RESEARCH STRATEGY

Significance

Health information technology is expected to play a key role in improving the quality of U.S. health care. Electronic medical records (EMRs) have been promoted as a means to improve the quality of medical care in the U.S. Encouraging the adoption of electronic medical records in primary care settings has been an objective of federal health policy. Through passage of the Health Information Technology for Economic and Clinical Health (HITECH) Act in 2009, Congress has reaffirmed the objective of increasing adoption of ambulatory care EMRs. Federal incentives linked to the adoption of EMR systems require physicians and hospitals to demonstrate “meaningful use” of certified EMR technology. “Meaningful use” refers to the use of EMRs not only for gathering and storing information, but also for tracking and improving specific outcomes. Underlying these policy developments is the assumption that EMR use, if meaningful, will improve the quality, safety, and efficiency of patient care.

Studies on the effect of EMR use on quality of care have found mixed evidence of benefits. There is evidence that while use of EMR as a binary factor is not associated with higher-quality care, the use of specific EMR features is associated with better performance on certain quality measures, suggesting that the extent of use of an EMR is important for quality improvement. For example, implementation of computerized clinical decision-support systems has been shown to improve clinician performance in specific settings. In particular, a number of studies have demonstrated that EMRs can
decrease the incidence of medication errors and adverse drug events\textsuperscript{13-19}. While such positive findings are on the rise, other studies of typical EMR use have found no effect or negative effect\textsuperscript{20-24}. A possible explanation for this EMR use-quality gap is that even among established users of EMR technology, effective use is not widespread. In a study that evaluated outcomes of actionable reminders, many intervention physicians reported not using or even not being aware of the new features of the EMR system\textsuperscript{25}. In another study, physicians who predominantly dictated their notes performed worse in EMR-assessed quality measures than physicians who more intensively interacted with EMR through the use of structured documentation\textsuperscript{26}. Overall, the inconsistent findings suggest that access to EMR technology does not guarantee its meaningful use\textsuperscript{27}, and provider education and training are needed to fully realize its potential benefits.

Documentation of clinical information, the process in which clinical events and provider-patient interactions are translated into electronic data, is the most critical step in ensuring that the recorded information is meaningful for quality monitoring and improvement. However, little empirical research has focused on evaluating how specific characteristics of electronic documentation reflect the care delivered or how they relate to clinical outcomes. In particular, very little is known about how characteristics of narrative EMR documentation reflect the quality of care delivered to the patient.

This project aimed at addressing scientific knowledge gaps regarding how narrative electronic documentation may reflect important components of provider-patient encounter that are associated with better patient outcomes. Previously published studies
indicated, for example, that only distinct – but not copy-pasted – documentation of lifestyle counseling was associated with faster achievement of blood glucose control by patients with diabetes\textsuperscript{28}. In this project we sought to further augment our understanding of the relationship between characteristics of narrative EMR documentation and patient outcomes by developing metrics of narrative documentation that may more specifically reflect the underlying quality of patient care. To this end we evaluated the relationship between several quantitative characteristics of narrative EMR documentation and patient outcomes on the example of blood glucose control in patients with diabetes.

Chronic hyperglycemia is the hallmark metabolic abnormality associated with diabetes mellitus\textsuperscript{29}. Achieving specific glycemic goals can substantially reduce morbidity and diabetes-specific complications including retinopathy, nephropathy, periodontal disease, and cardiovascular disease complications\textsuperscript{29-31}. The most recent glycemic goal recommended by the American Diabetes Association is a hemoglobin A1c level less than 7\%\textsuperscript{29}. While clinical trials convincingly demonstrate the salutary effects of good glycemic control, a large majority of patients with DM do not reach evidence-based glycemic goals\textsuperscript{32}.

Over-nutrition and a sedentary lifestyle with consequent weight gain are the major factors that increase glucose levels in patients with diabetes\textsuperscript{29}. Lifestyle interventions that address these factors have been demonstrated to have beneficial effects on glycemic control\textsuperscript{33-35}. It has been shown that weight reduction and an increase in daily energy expenditure decrease insulin resistance and increase glucose tolerance\textsuperscript{36}. A meta-analysis
has shown that exercise significantly improves glycemic control in people with type 2 diabetes even without weight loss\textsuperscript{37}. Avoiding being overweight and exercising regularly may also provide a protective effect against hypertension and hyperlipidemia, subsequently decreasing morbidity and mortality from cardiovascular disease\textsuperscript{36}. Consequently, lifestyle intervention with diet, exercise, and weight loss counseling is recommended in national and international guidelines for managing type 2 diabetes\textsuperscript{39,38}. In a recent study, Morrison \textit{et al.} investigated the long-term effects of lifestyle counseling on patients with DM in routine clinical settings\textsuperscript{39}. In this retrospective cohort study of over 30,000 patients in a 10-year study period, lifestyle counseling was strongly associated with faster achievement of A1c, blood pressure, and LDL cholesterol control, confirming the findings of controlled clinical trials.

The present study adds to our current scientific knowledge in several important ways. First, the findings of this project promote a better understanding of factors associated with improved patient outcomes in routine diabetes care. Enhancing our knowledge on factors associated with better managing this devastating disease benefit millions of patients in the U.S. and potentially worldwide.

Second, by identifying characteristics of electronic documentation associated with improved patient outcomes, this project contributes to developing novel types of surveillance mechanisms to monitor provider performance and quality of patient care. Meaningful Use incentive program has quality measurement as one important EMR capability. Our study provides insights into the relationship between electronic

16
documentation and quality of care, and thus will promote meaningful use of EMR technology.

Third, this project has developed a generalizable approach to studying the relationship between quantitative characteristics of narrative provider notes and quality of patient care. The approach we developed lays down the foundation for further applications of similar technology not only in conducting clinical research but also in developing novel features of EMR systems for monitoring quality and supplying feedback to providers. Since the tools and methods we developed should be readily generalizable to other health care settings, the benefits of this study extend beyond the current study population.

**Innovation**

We leveraged the extensive EMR infrastructure of our research data repository to develop an innovative technology — a suite of software tools that permit cost-and time-efficient computational analysis of narrative EMR data from a large number of patient encounters. This technology was validated and is generalizable to other healthcare settings.

Free-text physician notes are a significant component of an EMR system\(^{40,41}\). Narrative provider notes contain highly granular information regarding a patient’s care, status, and outcomes\(^{42}\). Unlike structured data where certain information details can be lost, narrative text provides contextual information to help clinicians to more thoroughly
understand the documented clinical events. In addition, some clinical information only exists in unstructured format. For example, in a study that evaluated home blood pressure documentation, it was shown that home blood pressure measurements were nearly always recorded in narrative notes rather than in the structured data\textsuperscript{43}. In a large-scale analysis of narrative vs. structured documentation of anti-hypertensive medication intensification, Turchin et al. found that medication intensification events were frequently only documented in narrative provider notes\textsuperscript{44}. Another study found that most adverse reactions to statins documented in narrative provider notes were not recorded in structured data\textsuperscript{45}.

Since large amounts of valuable clinical data are stored as free text, there is a need to develop automated methods to extract focused clinical information from narrative EMR data. Natural language processing (NLP) is a technology that enables rapid extraction of specific clinical information through computational analysis of clinical narratives\textsuperscript{46}. NLP offers a powerful alternative to either limited administrate data or labor-intensive, expensive manual chart reviews. NLP approach has been successfully used in clinical research, including identification of cohorts of patients with specific clinical characteristics, inactive medications, medication non-adherence, and adverse drug reactions, among many other applications\textsuperscript{44,47-59}. This project utilized a previously validated NLP system to exhaustively extract all sentences documenting lifestyle counseling from each narrative provider note.
Accurate documentation of clinical information is essential for ensuring that the recorded information can be successfully used for quality monitoring and improvement. One of the consequences of deploying EMR is growth in copying and pasting of clinical notes\textsuperscript{60-62}. Studies showed that 9 to 25 percent of electronic provider notes contained text fragments copied from elsewhere in the EMR system\textsuperscript{63-65}. While copying itself does not increase the probability of documentation error, there are significant risks associated with inappropriate copy and paste, including introduction of inconsistencies in the record and error propagation\textsuperscript{61,63}. Turchin et al. evaluated the relationship between copied lifestyle counseling documentation and blood glucose control in patients with DM\textsuperscript{28}. This large-scale retrospective study showed that, unlike original records, copied documentation of lifestyle counseling was not associated with faster achievement of target A1c level, and its effect was indistinguishable from no counseling at all. The findings of this study suggest that copied electronic documentation may not be a reliable representation of the actual care delivered.

Based upon the findings of Turchin et al.’s study on copy-paste documentation, we hypothesized that more diverse form of documentation is associated with better patient outcomes. In order to test this hypothesis, we developed an algorithm to quantify the heterogeneity of documentation by using the Levenshtein distance. Levenshtein distance, defined as the total number of edit operations (insertions, deletions, and substitutions) necessary to convert one string into another, is a metric that represents how similar two strings are\textsuperscript{66}. This metric was previously used in clinical informatics research.
in the identification of misspelled words and quantification of clinical narrative redundancy\textsuperscript{53,67}. To our knowledge, this is the first study that applied the concept of Levenshtein distance to evaluating the relationship between characteristics of narrative EMR documentation and patient outcomes.

**Approach**

**Design**

We conducted a retrospective cohort study to investigate whether quantitative characteristics of narrative EMR documentation of lifestyle counseling are associated with blood glucose control in patients with diabetes.

**Data**

We used patient data from Partners HealthCare System, an integrated healthcare delivery network that comprises several academic and community hospitals and private physician groups in eastern Massachusetts, including the founding members Brigham and Women’s Hospital (BWH) and Massachusetts General Hospital (MGH). The network serves a patient population of approximately 1.3 million, including an estimated diabetes population of 100,000. The network’s primary care physicians provide the majority of its diabetes-related care. The main EMR used in Partners HealthCare ambulatory clinics is the Longitudinal Medical Record (LMR). The LMR is an ONC-ATCB (Meaningful Use) certified complete electronic medical record system, developed internally at Partners HealthCare. LMR allows structured entry of data elements (e.g. medication lists, allergies,
problem lists) as well as narrative text (e.g. progress notes, radiology and pathology reports). LMR has a number of decision support features but does not include any decision support for lifestyle counseling. 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.

Demographic information, laboratory data, and the text of physician notes were obtained from the Research Partners Data Registry (RPDR). RPDR is a data warehouse that serves as a central clinical data repository for researchers at participating hospitals and clinics within the Partners HealthCare system. The following demographic data were obtained for each study patient: age, gender, race / ethnicity, primary language, health insurance, and median income by zip code.

**Study Cohort**

Patients with diabetes mellitus followed by primary care physicians (PCPs) affiliated with BWH and MGH between 01/01/2000 and 01/01/2011 were studied. Patients were included in the analysis if they are at least 18 years old, had a documented diagnosis of diabetes mellitus or hemoglobin A1c ≥7.0%, and had been followed by a physician in a primary care specialty for at least two years during the study period. Diagnosis of diabetes was ascertained by analyzing the text of physician notes in the EMR as previously described\(^{51}\). Patients were excluded from the analysis if their zip code was missing (to enable adjustment for median household income by zip code).
The Institutional Review Board at Partners HealthCare System approved the study and the requirement for written informed consent was waived.

**Study Measurements**

A single *hyperglycemic period* served as the unit of analysis for the evaluation of outcomes in all Specific Aims. A hyperglycemic period started at the first available hemoglobin A1c measurement of 7.0% or greater and ended at the first A1c measurement lower than 7.0% or at the end of the study period if A1c never reached treatment target (Figure 1). We used <7.0% as the recommended treatment goal of hemoglobin A1c level in accordance with the current published guidelines\textsuperscript{29}. Each patient may have contributed multiple periods if measures fluctuated above and below the target A1c level during the study period. If multiple A1c measurements were made on the same day, the lowest value was used for the analysis. Transient elevations of hemoglobin A1c, defined as a single elevated measurement followed by a fall below the target level at the next measurement with no interceding anti-hyperglycemic treatment intensification, were excluded from the analysis. Hyperglycemic periods without at least one annual encounter with BWH/MGH primary care physician were excluded to eliminate patients not actively treated in these practices. Hyperglycemic periods without any medication information available in the EMR were excluded to enable inclusion of insulin treatment as a confounder variable. Periods that contained more than one encounter with an endocrinologist or a diabetologist that addressed diabetes, as ascertained using a combination of billing data and computational analysis of the notes, were excluded in order to ensure a single source of
diabetes care. Physician specialty was identified using a combination of the information available from the Massachusetts Board of Registration in Medicine and the specialty of the clinic where the physician practiced. Hyperglycemic periods where rate of change of A1c is greater than 3 SD from the mean were excluded from the analysis to eliminate likely measurement errors.

**Figure 1.** Illustration of a hyperglycemic period. A hyperglycemic period started at the first available hemoglobin A1c measurement of 7.0% or greater and ended at the first A1c measurement lower than 7.0% or at the end of the study period if A1c never reached treatment target.

Time to glycemic control for any hyperglycemic period was defined as the length of that hyperglycemic period. Frequency of lifestyle counseling was calculated as the average monthly number of encounters with documented lifestyle counseling. Medication intensification rate was defined as the average monthly number of episodes of anti-
hyperglycemic treatment 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 electronic provider notes as previously validated

Sentences dedicated to documenting lifestyle counseling were extracted automatically from narrative provider notes using a previously validated natural language processing system with sensitivity of 91.4% and specificity of 88.2%\textsuperscript{28}. The counseling sentences were categorized into 3 counseling categories (diet, exercise, and weight loss).

The NLP software was implemented in Perl 5.16.3.

Calculation of Documentation Heterogeneity

Documentation heterogeneity was calculated for two consecutive notes with documented lifestyle counseling as the Levenshtein desistance between the relevant sentences, normalized by the length of the longer sentence. In this study, we used the Levenshtein distance algorithm that permitted transposition of two adjacent letters (Damerau-Levenshtein distance).\textsuperscript{68}

The algorithm for calculation of documentation heterogeneity is illustrated in Figure 2, and an example of calculation procedures is shown in Figure 3. The normalized Levenshtein distance (NLD) between two sentences was calculated by dividing the Levenshtein distance by the length (in characters) of the longer sentence. In calculating the NLD between lifestyle counseling notes \(i\) and \(i+1\), we ignored any notes between \(i\) and \(i+1\) that did not have lifestyle counseling documented. When comparing two notes
we first calculated NLDs between all sentences documenting lifestyle counseling in each of the notes. For example, if note $i$ had 3 sentences and note $i+1$ had 4 sentences, we calculated a total of 12 NLDs. For each of the sentences $[S_{1,i+1}, S_{2,i+1}, ..., S_{m,i+1}]$ in note $i+1$, we calculated inter-sentence NLD to each of the sentences $[S_{1,i}, S_{2,i}, ..., S_{m,i}]$ in note $i$. We repeated this process for all sentences in note $i+1$. To obtain the inter-note NLD between note $i$ and note $i+1$, we took the highest of the inter-sentence NLDs calculated between note $i$ and note $i+1$. We then calculated the mean inter-note NLD for all notes documenting lifestyle counseling during the hyperglycemic period. These comparisons were made separately within each of the 3 categories of counseling: diet mean inter-note NLD, exercise mean inter-note NLD, and weight loss mean inter-note NLD. Finally, we calculated the mean of these three NLDs to obtain the documentation heterogeneity for the hyperglycemic period. In order to be included in the analysis, the hyperglycemic periods did not need to have all the three categories of lifestyle counseling. When the mean inter-note NLD was missing for one or two counseling categories, documentation heterogeneity for the period was calculated based upon the counseling categories for which the mean inter-note NLD was available.
Figure 2. Algorithm for calculation of documentation heterogeneity. Hyperglycemic period $q$ contains notes $[1, \ldots, p]$. Note $i$ contains $[S_{1,i}, S_{2,i}, \ldots, S_{m,i}]$. For each of sentences $[S_{1,i+1}, S_{2,i+1}, \ldots, S_{m,i+1}]$ in note $i+1$, we calculated the inter-sentence $L$ ($L$: normalized Levenshtein distance) to each of sentences $[S_{1,i}, S_{2,i}, \ldots, S_{m,i}]$ in note $i$. To obtain the inter-note $L$ between note $i$ and note $i+1$, we took the highest of the inter-sentence $L$’s (indicated in the matrix in the figure). To obtain the mean inter-note $L$ for hyperglycemic period $q$, we calculated the mean of the $p-1$ inter-note $L$’s. These operations were performed for each counseling category, and subsequently the mean of the $3$ categories was calculated to obtain the documentation heterogeneity of the hyperglycemic period.
Figure 3. Example of documentation heterogeneity calculation. The above hyperglycemic contains notes [1, 2, 3]. Note 1 contains sentences \([S_{1,1}, S_{2,1}, S_{3,1}]\), note 2 contains sentences \([S_{1,2}, S_{2,2}, S_{3,2}, S_{4,2}]\), and note 3 contains sentences \([S_{1,3}, S_{2,3}]\). For each of sentences \([S_{1,2}, S_{2,2}, S_{3,2}, S_{4,2}]\) in note 2, the inter-sentence \(L\) (\(L\): normalized Levenshtein distance) to each of sentences \([S_{1,1}, S_{2,1}, S_{3,1}]\) in note 1 is calculated. The highest of the 12 inter-sentence \(L\)'s (indicated in the first matrix in the figure) represent the inter-note \(L\) between note 1 and note 2. These operations are repeated between note 2 and note 3 (the results are shown in the second matrix in the figure). The mean of the 2 inter-note \(L\)'s represent the mean inter-note \(L\) for this hyperglycemic period. Note that all the sentences in this example belong to the same counseling category.
Threshold sentence length determination

To ensure comparisons between single sentences in calculating NLD, sentences longer than the threshold length were excluded from the output of the NLP system. Manual review of randomly selected 600 sentences by one reviewer determined that sentences longer than 100 characters were unlikely to represent a single lifestyle counseling sentence. After exclusion of sentences longer than 100 characters, 90% of the sentences were single sentences documenting lifestyle counseling. This threshold length was validated by manual review of another 600 randomly selected sentences. After exclusion of sentences longer than 100 characters, 91% of the sentences were single sentences documenting lifestyle counseling.

Exclusion of copied/duplicate counseling sentences

Since Turchin et al.’s study has shown that copied documentation of lifestyle counseling is not associated with improvement in glucose control, we performed the following procedures to exclude copied/duplicate counseling sentences.

After calculation of the mean inter-note NLD for each counseling category, we excluded hyperglycemic periods in which 1) the mean inter-note NLD was zero for all 3 categories; or 2) the mean inter-note NLD was zero for 1 or 2 categories and incalculable for the rest (inter-note NLD is incalculable if there are less than 2 notes per counseling category). In addition, if the mean inter-note NLD was zero for any counseling category, the zero value was ignored in calculating the mean inter-note NLD of the 3 counseling
categories, in order to exclude counseling category in which the counseling documentation consisted entirely of copied sentences throughout the hyperglycemic period.

**Calculation of documentation intensity**

The algorithm for calculation of documentation intensity is illustrated in Figure 4, and an example of calculation procedures is shown in Figure 5. Sentences from all the 3 counseling categories were included. We calculated the mean number of characters per note dedicated to documenting lifestyle counseling. Only distinct counseling sentences were included (i.e. if the same sentences appeared multiple times during the period, they were counted only once). 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 be able to assess relative contribution of documentation heterogeneity and documentation intensity.

Sample data and calculation results of documentation heterogeneity and documentation intensity are shown in Figure 6. The software for calculation of documentation heterogeneity and documentation intensity was implemented in Perl 5.16.13. The overall study design is schematically represented in Figure 7.
Figure 4. Algorithm for calculation of documentation intensity. Hyperglycemic period \( q \) contains notes \([1, \ldots, p]\) that document lifestyle counseling (notes with no documented lifestyle counseling were ignored). Note \( i \) contains sentences \([S_{1,i}, S_{2,i}, \ldots, S_{m,i}]\) dedicated to documenting lifestyle counseling. Documentation intensity for the hyperglycemic period was calculated as the sum of the lengths (in characters) of sentences \([S_{1,i}, S_{2,i}, \ldots, S_{m,i}]\) for notes \([1, \ldots, p]\), divided by the number of notes \( p \).
Figure 5. Example of documentation intensity calculation. This hyperglycemic contains notes [1, 2, 3] that document lifestyle counseling. Note 1 contains sentences \([S_{1,1}, S_{2,1}, S_{3,1}]\), note 2 contains sentences \([S_{1,2}, S_{2,2}, S_{3,2}, S_{4,2}]\), and note 3 contains sentences \([S_{1,3}, S_{2,3}]\), which are dedicated to documenting lifestyle counseling. Documentation intensity is calculated as the sum of the sentence lengths divided by the number of notes, in this case 3.
Figure 6. Sample data and calculation results of documentation heterogeneity and documentation intensity. *Documentation heterogeneity represented by the normalized Levenshtein distance.

Statistical Analysis

Summary statistics were calculated by using frequencies and proportions for categorical data, and by using means, standard deviations, and medians for continuous variables.

In univariate analysis, for ease of interpretation we divided the hyperglycemic periods into tertiles for each predictor variable. The log-rank test was used to compare Kaplan-Meier cumulative incidence curves for time to achievement of A1c target.
Marginal Cox proportional-hazards regression models for clustered data were constructed to estimate the effect of predictor variables on time to achievement of A1c target while adjusting for covariates and accounting for clustering within individual patients. The covariates used in the Cox model included patient age, gender, race/ethnicity, primary language, health insurance, income, initial A1c level, frequency of A1c measurement, treatment with insulin, frequency of medication intensification, frequency of encounters with documented lifestyle counseling, Charlson comorbidity index, and presence of obesity during the period. We also performed an analysis accounting for clustering within individual providers. Primary care provider was determined for each hypoglycemic period as the provider who contributed the highest number of notes used for calculation of documentation heterogeneity and documentation intensity for the period. Simes-Hochberg correction was used to adjust the p-values for multiple comparisons.

In addition, multiple linear regression analyses were performed in order to determine the relationship between the confounder variables and the predictor variables.

As a sensitivity analysis, we performed a multivariable analysis using hypoglycemic periods that included four categories of previously excluded periods: 1) no income information; 2) no medication information; 3) transient elevation of A1c; and 4) two or more encounters with endocrinologists. For this analysis, we included two additional covariates for provider characteristics: 1) provider age; and 2) provider gender.
All analyses were performed with SAS statistical software (version 9.3, SAS Institute, Cary, NC).

Figure 7. Schematic representation of the study protocol.

RESULTS

We identified 37,141 hyperglycemic periods meeting the initial inclusion criteria, contributed by adult patients with diabetes who were followed by BWH- or MGH-affiliated PCP for at least 2 years during the study period, with at least 1 encounter per year during the hyperglycemic period. We excluded 1,802 hyperglycemic periods contributed by patients missing the income information, 3,844 hyperglycemic periods in
which patients were treated by an endocrinologist or a diabetologist, 1,990 hyperglycemic periods with no medication records, 1,317 hyperglycemic periods with suspected A1c measurement errors, and 5,997 hyperglycemic periods where only transient elevations of A1c were recorded. We further excluded 1,788 hyperglycemic periods with no documented lifestyle counseling, 2,119 hyperglycemic periods which contained no counseling notes with sentences meeting the threshold length criteria, 4,266 hyperglycemic periods which contained less than 2 notes with documented lifestyle counseling per counseling category for all 3 counseling categories, and 424 hyperglycemic periods in which lifestyle counseling was documented entirely by copied sentences throughout the hyperglycemic period in each counseling category. The remaining 13,594 hyperglycemic periods were included in the analysis. Counts of hyperglycemic periods excluded from the analysis are summarized in Figure 8.
Figure 8. Counts of hyperglycemic periods excluded from the analysis.

A total of 10,870 unique patients contributed to the 13,594 hyperglycemic periods. The mean age of the study patients was 59 years; about half of the patients were women, and the majority of patients indicated English as their primary language. The mean
follow-up time during the study period was 6.8 years. Their A1c level was above the recommended target for >50% of that time. Characteristics of the study patients are summarized in Table 1.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)*</td>
<td>59.3</td>
</tr>
<tr>
<td>Women, n (%)</td>
<td>5,635</td>
</tr>
<tr>
<td>Race/ethnicity, n (%)</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>6,447</td>
</tr>
<tr>
<td>Black</td>
<td>1,576</td>
</tr>
<tr>
<td>Hispanic</td>
<td>1,772</td>
</tr>
<tr>
<td>Asian</td>
<td>370</td>
</tr>
<tr>
<td>Other†</td>
<td>705</td>
</tr>
<tr>
<td>English as the primary language, n (%)</td>
<td>8,623</td>
</tr>
<tr>
<td>Health insurance, n (%)</td>
<td></td>
</tr>
<tr>
<td>Private</td>
<td>4,336</td>
</tr>
<tr>
<td>Medicare</td>
<td>5,162</td>
</tr>
<tr>
<td>Medicaid</td>
<td>1,183</td>
</tr>
<tr>
<td>None/unknown</td>
<td>189</td>
</tr>
<tr>
<td>Median income by zip code, mean ($1000s)</td>
<td>50.3</td>
</tr>
<tr>
<td>No. of uncontrolled periods</td>
<td>1.3</td>
</tr>
<tr>
<td>Hemoglobin A1c (%)</td>
<td>8.3</td>
</tr>
<tr>
<td>BMI, mean (SD, % patients with measures)</td>
<td>32.9</td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>6.4</td>
</tr>
<tr>
<td>Follow-up time (months)</td>
<td>82.2</td>
</tr>
<tr>
<td>Total time above treatment target (months)</td>
<td>41.5</td>
</tr>
</tbody>
</table>

Table 1. Patient Characteristics. Data are mean (SD), unless otherwise indicated. *Age calculated at the start date of the first hyperglycemic period. †Includes unknown.
The included hyperglycemic periods had a mean length of 33 months. The mean initial hemoglobin A1c level was 8.3%, and the mean rate of anti-hyperglycemic medication intensification was approximately once a year. In about 30% of the hyperglycemic periods patients were on insulin. Lifestyle counseling was documented at the mean rate of once every 3.3 months. The mean rate of A1c testing was once in 3.6 months. Characteristics of the included hyperglycemic periods are summarized in Table 2.

<table>
<thead>
<tr>
<th>Study periods, $n$</th>
<th>13,594</th>
</tr>
</thead>
<tbody>
<tr>
<td>Period length (months)</td>
<td>33.2 (28.1)</td>
</tr>
<tr>
<td>Mean initial hemoglobin A1c (%)</td>
<td>8.3 (1.6)</td>
</tr>
<tr>
<td>Periods where treatment target was reached, $n$ (%)</td>
<td>8,300 (61.1)</td>
</tr>
<tr>
<td>Rate of A1c measurement per month</td>
<td>0.28 (0.16)</td>
</tr>
<tr>
<td>Rate of lifestyle counseling per month</td>
<td>0.30 (0.24)</td>
</tr>
<tr>
<td>Rate of medication intensification per month</td>
<td>0.10 (0.13)</td>
</tr>
<tr>
<td>Periods with patients on insulin, $n$ (%)</td>
<td>4,118 (30.3)</td>
</tr>
<tr>
<td>Periods with patients who are obese (BMI &gt; 30), $n$ (%)</td>
<td>9,088 (66.9)</td>
</tr>
<tr>
<td>Documentation heterogeneity*</td>
<td>0.69 (0.15)</td>
</tr>
<tr>
<td>Documentation intensity (characters(note)</td>
<td>90.8 (44.6)</td>
</tr>
</tbody>
</table>

Table 2. Hyperglycemic period characteristics. Data are mean (SD), unless otherwise indicated. *Documentation heterogeneity represented by the normalized Levenshtein distance.

A total of 183,611 sentences from 92,671 provider notes were analyzed to calculate the documentation heterogeneity and documentation intensity. The mean
documentation heterogeneity, represented by the normalized Levenshtein distance, was 0.7. The mean documentation intensity was 90.8 characters per note.

In univariate analysis, median time to achievement of target A1c level rose progressively both at the lower documentation heterogeneity and at the lower documentation intensity. Comparing hyperglycemic periods in the highest versus lowest tertile by documentation heterogeneity and documentation intensity, median time to A1c < 7.0% was 26 versus 39 months, and 24 versus 39 months, respectively (P < 0.0001 for all). Kaplan-Meier cumulative incidence curves for tertiles by documentation heterogeneity and intensity are shown in Figure 9.
Figure 9. Documentation characteristics and time to A1c target. Kaplan-Meier curves for time to A1c target were plotted for tertiles by documentation heterogeneity and intensity. Distinct hyperglycemic periods for the same patient were analyzed separately. A: Documentation heterogeneity and time to A1c target (*Heterogeneity represented by the normalized Levenshtein distance; P < 0.0001 by log-rank test). B. Documentation intensity and time to A1c target (P < 0.0001 by log-rank test).

In multivariable Cox proportional hazards models adjusted for the patients’ demographic characteristics, initial A1c level, frequency of A1c measurement, treatment with insulin, frequency of medication intensification, frequency of encounters with documented lifestyle counseling, Charlson comorbidity index, presence of obesity during the period, and clustering within individual patients, an increase of documentation
heterogeneity by 1 unit and an increase of documentation intensity by 100 characters/note was associated with hazard ratios of 1.68 (95% CI 1.33 to 2.12; P < 0.0001) and 1.70 (95% CI 1.59 to 1.81; P < 0.0001) for time to achievement of A1c target, respectively. Effects of patient, treatment, and documentation characteristics on time to A1c control are shown in Table 3.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio</th>
<th>95% Confidence limits</th>
<th>P value ($\chi^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.923</td>
<td>0.873</td>
<td>0.975</td>
</tr>
<tr>
<td>White race</td>
<td>0.974</td>
<td>0.916</td>
<td>1.035</td>
</tr>
<tr>
<td>English speaker</td>
<td>1.020</td>
<td>0.954</td>
<td>1.090</td>
</tr>
<tr>
<td>Income, per $1000 increase</td>
<td>0.998</td>
<td>0.997</td>
<td>1.000</td>
</tr>
<tr>
<td>Government Insurance</td>
<td>0.975</td>
<td>0.916</td>
<td>1.037</td>
</tr>
<tr>
<td>On insulin</td>
<td>0.380</td>
<td>0.352</td>
<td>0.410</td>
</tr>
<tr>
<td>Obesity during the period</td>
<td>0.987</td>
<td>0.924</td>
<td>1.054</td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>1.012</td>
<td>1.005</td>
<td>1.020</td>
</tr>
<tr>
<td>Rate of A1c testing, per 3 month</td>
<td>3.579</td>
<td>2.942</td>
<td>4.354</td>
</tr>
<tr>
<td>Rate of antihyperglycemic medication intensification, per month</td>
<td>3.389</td>
<td>2.186</td>
<td>5.255</td>
</tr>
<tr>
<td>Rate of lifestyle counseling, per month</td>
<td>7.697</td>
<td>5.925</td>
<td>10.000</td>
</tr>
<tr>
<td>A1c level at start of period</td>
<td>0.849</td>
<td>0.831</td>
<td>0.863</td>
</tr>
<tr>
<td>Age (10 years)</td>
<td>1.065</td>
<td>1.036</td>
<td>1.095</td>
</tr>
<tr>
<td>Documentation heterogeneity*</td>
<td>1.680</td>
<td>1.331</td>
<td>2.120</td>
</tr>
<tr>
<td>Documentation intensity (100 characters/note)</td>
<td>1.698</td>
<td>1.592</td>
<td>1.811</td>
</tr>
</tbody>
</table>

Table 3. Effects of patient, treatment, and documentation characteristics on time to A1c control accounting for clustering within individual patients. *Documentation heterogeneity represented by the normalized Levenshtein distance. †Significant after Simes-Hochberg correction.
A total of 1,211 unique primary care providers were identified for the included hyperglycemic periods. The relationship of documentation heterogeneity and intensity with time to A1c control did not change substantially when the multivariable analysis was adjusted for clustering within individual providers (Table 4).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio</th>
<th>95% Confidence limits</th>
<th>P value ($\chi^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.922</td>
<td>0.857</td>
<td>0.992</td>
</tr>
<tr>
<td>White race</td>
<td>0.978</td>
<td>0.906</td>
<td>1.056</td>
</tr>
<tr>
<td>English speaker</td>
<td>1.026</td>
<td>0.945</td>
<td>1.114</td>
</tr>
<tr>
<td>Income, per $1000 increase</td>
<td>0.999</td>
<td>0.997</td>
<td>1.001</td>
</tr>
<tr>
<td>Government Insurance</td>
<td>0.975</td>
<td>0.905</td>
<td>1.051</td>
</tr>
<tr>
<td>On insulin</td>
<td>0.352</td>
<td>0.323</td>
<td>0.383</td>
</tr>
<tr>
<td>Obesity during the period</td>
<td>0.968</td>
<td>0.909</td>
<td>1.031</td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>1.014</td>
<td>1.004</td>
<td>1.023</td>
</tr>
<tr>
<td>Rate of A1c testing, per 3 month</td>
<td>3.445</td>
<td>2.563</td>
<td>4.630</td>
</tr>
<tr>
<td>Rate of antihyperglycemic medication intensification, per month</td>
<td>3.715</td>
<td>2.437</td>
<td>5.661</td>
</tr>
<tr>
<td>Rate of lifestyle counseling, per month</td>
<td>8.292</td>
<td>6.007</td>
<td>11.446</td>
</tr>
<tr>
<td>A1c level at start of period</td>
<td>0.857</td>
<td>0.840</td>
<td>0.875</td>
</tr>
<tr>
<td>Age (10 years)</td>
<td>1.058</td>
<td>1.022</td>
<td>1.096</td>
</tr>
<tr>
<td>Documentation heterogeneity*</td>
<td>1.858</td>
<td>1.276</td>
<td>2.706</td>
</tr>
<tr>
<td>Documentation intensity (100 characters/note)</td>
<td>1.772</td>
<td>1.617</td>
<td>1.941</td>
</tr>
</tbody>
</table>

Table 4. Effects of patient, treatment, and documentation characteristics on time to A1c control accounting for clustering within individual providers. *Documentation heterogeneity represented by the normalized Levenshtein distance. †Significant after Simes-Hochberg correction.
Correlation analyses with multiple linear regression models showed that there were only limited correlations between the confounder variables and the predictor variables. The overall model fit was $R^2 = 0.02$ for documentation heterogeneity and $R^2 = 0.06$ for documentation intensity (Table 5 and Table 6).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter estimate</th>
<th>95% Confidence limits</th>
<th>$P$ value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.6960</td>
<td>0.6732 - 0.7188</td>
<td>$&lt; 0.0001$†</td>
</tr>
<tr>
<td>Female</td>
<td>0.0007</td>
<td>-0.0043 - 0.0057</td>
<td>0.7873</td>
</tr>
<tr>
<td>White race</td>
<td>-0.0155</td>
<td>-0.0215 - -0.0094</td>
<td>$&lt; 0.0001$</td>
</tr>
<tr>
<td>English speaker</td>
<td>0.0089</td>
<td>0.0020 - 0.0158</td>
<td>0.0113†</td>
</tr>
<tr>
<td>Income, per $1000$ increase</td>
<td>0.0002</td>
<td>0.0000 - 0.0003</td>
<td>0.0323†</td>
</tr>
<tr>
<td>Government Insurance</td>
<td>-0.0005</td>
<td>-0.0066 - 0.0056</td>
<td>0.8779</td>
</tr>
<tr>
<td>On insulin</td>
<td>-0.0163</td>
<td>-0.221 - -0.0011</td>
<td>$&lt; 0.0001$†</td>
</tr>
<tr>
<td>Obesity during the period</td>
<td>0.0082</td>
<td>0.0029 - 0.0136</td>
<td>0.0025†</td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>0.0006</td>
<td>-0.0001 - 0.0012</td>
<td>0.0788</td>
</tr>
<tr>
<td>Rate of A1c testing, per 3 month</td>
<td>0.0345</td>
<td>0.002/ - 0.041/</td>
<td>$&lt; 0.0001$†</td>
</tr>
<tr>
<td>Rate of antihyperglycemic medication intensification, per month</td>
<td>0.0664</td>
<td>0.0440 - 0.0889</td>
<td>$&lt; 0.0001$†</td>
</tr>
<tr>
<td>Rate of lifestyle counseling, per month</td>
<td>-0.0974</td>
<td>-0.1114 - -0.0834</td>
<td>$&lt; 0.0001$†</td>
</tr>
<tr>
<td>A1c level at start of period</td>
<td>0.0016</td>
<td>0.0000 - 0.0032</td>
<td>0.0590</td>
</tr>
<tr>
<td>Age (10 years)</td>
<td>-0.0089</td>
<td>-0.0008 - -0.0031</td>
<td>$&lt; 0.0001$†</td>
</tr>
</tbody>
</table>

**Table 5.** Correlations between the confounder variables and documentation heterogeneity.

†Significant at $P < 0.05$. 
<table>
<thead>
<tr>
<th>Variable</th>
<th>Parameter estimate</th>
<th>95% Confidence limits</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>0.9641</td>
<td>0.8965 - 1.0318</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Female</td>
<td>0.0166</td>
<td>0.0017 - 0.0315</td>
<td>0.0289</td>
</tr>
<tr>
<td>White race</td>
<td>-0.0116</td>
<td>-0.0294 - 0.0062</td>
<td>0.2013</td>
</tr>
<tr>
<td>English speaker</td>
<td>0.0857</td>
<td>0.0652 - 0.1061</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Income, per $1000 increase</td>
<td>0.0012</td>
<td>0.0036 - 0.0068</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Government Insurance</td>
<td>-0.0139</td>
<td>-0.0320 - 0.0042</td>
<td>0.1329</td>
</tr>
<tr>
<td>On insulin</td>
<td>-0.1084</td>
<td>-0.1256 - -0.0913</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Obesity during the period</td>
<td>0.0523</td>
<td>0.0036 - 0.0681</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>-0.0045</td>
<td>-0.0064 - -0.0026</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Rate of A1c testing, per 3 month</td>
<td>0.1243</td>
<td>0.1030 - 0.1456</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Rate of antihyperglycemic medication intensification, per month</td>
<td>0.0719</td>
<td>0.0053 - 0.1384</td>
<td>0.0342</td>
</tr>
<tr>
<td>Rate of lifestyle counseling, per month</td>
<td>-0.0117</td>
<td>-0.1581 - -0.0751</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>A1c level at start of period</td>
<td>-0.0063</td>
<td>-0.0112 - -0.0015</td>
<td>0.0099</td>
</tr>
<tr>
<td>Age (10 years)</td>
<td>-0.0299</td>
<td>-0.0374 - -0.0224</td>
<td>&lt; 0.0001</td>
</tr>
</tbody>
</table>

**Table 6.** Correlations between the confounder variables and documentation intensity.

†Significant at P < 0.05.

A sensitivity analysis was performed for the secondary dataset that included previously excluded hyperglycemic periods (no income information, no medication information, treatment by endocrinologists, and transient elevation of A1c). The dataset included 19,562 hyperglycemic periods contributed by 14,863 unique patients. Provider age and provider gender were included as additional covariates. The results of the multivariable analysis showed that neither provider age nor provider gender had a significant relationship with time to A1c target. Inclusion of the previously excluded
periods and addition of the two provider characteristics covariates did not substantially change the relationship of documentation heterogeneity and intensity with time to A1c control (Table 7).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hazard ratio</th>
<th>95% Confidence limits</th>
<th>P value ($\chi^2$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Female</td>
<td>0.935</td>
<td>0.882 - 0.991</td>
<td>0.0238</td>
</tr>
<tr>
<td>White race</td>
<td>0.956</td>
<td>0.908 - 1.005</td>
<td>0.0801</td>
</tr>
<tr>
<td>English speaker</td>
<td>0.998</td>
<td>0.944 - 1.054</td>
<td>0.9314</td>
</tr>
<tr>
<td>Income, per $1000$ increase</td>
<td>0.998</td>
<td>0.997 - 0.999</td>
<td>0.0002 †</td>
</tr>
<tr>
<td>Government Insurance</td>
<td>0.920</td>
<td>0.868 - 0.975</td>
<td>0.0049 †</td>
</tr>
<tr>
<td>On insulin</td>
<td>0.386</td>
<td>0.346 - 0.410</td>
<td>&lt; 0.0001 †</td>
</tr>
<tr>
<td>Obesity during the period</td>
<td>0.970</td>
<td>0.922 - 1.020</td>
<td>0.2284</td>
</tr>
<tr>
<td>Charlson comorbidity index</td>
<td>1.015</td>
<td>1.008 - 1.021</td>
<td>&lt; 0.0001</td>
</tr>
<tr>
<td>Rate of A1c testing, per 3 month</td>
<td>3.435</td>
<td>2.688 - 4.390</td>
<td>&lt; 0.0001 †</td>
</tr>
<tr>
<td>Rate of antihyperglycemic medication intensification, per month</td>
<td>0.470</td>
<td>0.314 - 0.705</td>
<td>0.0003 †</td>
</tr>
<tr>
<td>Rate of lifestyle counseling, per month</td>
<td>6.520</td>
<td>4.631 - 9.180</td>
<td>&lt; 0.0001 †</td>
</tr>
<tr>
<td>A1c level at start of period</td>
<td>0.801</td>
<td>0.786 - 0.815</td>
<td>&lt; 0.0001 †</td>
</tr>
<tr>
<td>Age (10 years)</td>
<td>1.074</td>
<td>1.050 - 1.098</td>
<td>&lt; 0.0001 †</td>
</tr>
<tr>
<td>Provider female</td>
<td>0.992</td>
<td>0.945 - 1.041</td>
<td>0.7434</td>
</tr>
<tr>
<td>Provider age</td>
<td>0.981</td>
<td>0.958 - 1.005</td>
<td>0.1194</td>
</tr>
<tr>
<td>Documentation heterogeneity*</td>
<td>1.649</td>
<td>1.292 - 2.104</td>
<td>&lt; 0.0001 †</td>
</tr>
<tr>
<td>Documentation intensity (100 characters/note)</td>
<td>1.632</td>
<td>1.547 - 1.721</td>
<td>&lt; 0.0001 †</td>
</tr>
</tbody>
</table>

**Table 7.** Effects of patient, provider, treatment, and documentation characteristics on time to A1c control accounting for clustering within individual patients. *Documentation heterogeneity represented by the normalized Levenshtein distance. †Significant after Simes-Hochberg correction.
DISCUSSION

In this large, long-term retrospective study of documentation of lifestyle counseling in patients with diabetes, we have demonstrated that both higher documentation heterogeneity and higher documentation intensity are associated with faster achievement of glycemic control. This association was independent of other treatment processes including frequency of lifestyle counseling, medication intensification, and rate of A1c measurement. To our knowledge, this is the first study that has provided evidence that quantitative characteristics of narrative electronic documentation are associated with patient outcomes.

It is 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.\textsuperscript{69-72} While the relationship between the documentation characteristics and the actual content of lifestyle counseling provided is unknown, it is possible that the two measures of documentation characteristics we developed may be able to differentiate more intensive vs. less intensive lifestyle counseling. Providers who spend more time counseling the patient and provide more detailed instructions may also document the counseling episode in greater detail, which would be characterized by higher documentation heterogeneity and intensity. On the other hand, documentation of less detailed counseling session may be characterized by lower documentation heterogeneity and intensity. Further studies are needed to investigate, for example, whether counseling
episodes documented with higher documentation intensity is a reflection of more intense or more detailed lifestyle counseling.

The association between documentation characteristics and blood glucose control may be explained in the framework of the Chronic Care Model (CCM). The CCM was developed by Wagner and colleagues to help practices improve chronic illness outcomes through system changes that would make patient-centered, evidence-based care easier to accomplish. The CCM comprises 6 inter-related elements that are hypothesized to work together to promote patients’ self-efficacy and delivery of effective care: 1) health system, 2) self-management support, 3) delivery system design, 4) decision support, 5) clinical information systems, and 6) community resources and policies. The premise of the CCM is that successful implementation of these elements improve chronic illness outcomes through productive interactions between informed, activated patients and prepared, proactive practice teams. Studies provide substantial evidence that CCM approaches have been effective in managing diabetes in U.S. primary care settings.

One possible explanation for the association between documentation characteristics and blood glucose control is that the pathway to better clinical outcomes is mediated by improved patient self-care behavior (Figure 10). The two measures of documentation characteristics may reflect the quality of “productive interactions” between patients and providers in the CCM framework. Heterogeneity and intensity of lifestyle counseling documentation may be related to how “proactive” the provider is in “activating” the patients by “informing” them of the benefits of positive lifestyle changes.
and the importance of their role in managing the illness. Self-management support is the cornerstone of diabetes care in the CCM. More heterogeneous and more intense documentation of lifestyle counseling may reflect the presence of more effective self-management support for lifestyle changes. For example, discussion of more diverse aspects of diet changes may give more opportunities for patients to explore different ways to amend their diet. Better self-management support may in turn result in improved self-care behaviors, such as adoption of healthier diet and increased physical activities. Further studies are needed to evaluate how the documentation characteristics of lifestyle counseling are related to changes in patient behaviors.
Figure 10. Schematic representation of the hypothesized relationship between documentation characteristics and clinical outcomes in the Chronic Care Model (CCM) framework. Heterogeneity and intensity of documentation of lifestyle counseling may reflect the quality of “productive interactions” in the CCM framework. More heterogeneous and more intense documentation of lifestyle counseling may reflect the presence of more effective self-management support for lifestyle changes, which may in turn result in improved self-care behaviors, ultimately leading to improved clinical outcomes such as time to A1c control.

The findings of this study suggest that heterogeneity and intensity of documentation of lifestyle counseling in narrative provider notes may be used as indicators of quality of diabetes care. While further studies are needed to assess the generalizability of our findings, the approach we developed to quantify documentation characteristics will lay down the foundation for conducting similar studies to assess the relationship between narrative electronic documentation and quality of care.

Our study has several limitations. First, we could not assess the relationship between documentation characteristics and time to A1c control for the hyperglycemic periods that had only one note with documented lifestyle counseling per counseling category. These periods were excluded from the analyses as calculation of documentation heterogeneity by definition required two or more notes with documented lifestyle counseling per counseling category during the hyperglycemic period. Second, 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. Third, the software we used to identify documentation of lifestyle counseling did not provide details on the specific counseling approach (e.g. “Patient advised about diet.” vs. “Reviewed with the patient the importance of a low fat diet and a higher consumption of fruits, vegetables, and lean meats.”). Lastly, the retrospective nature of this study does not allow us to make causal inferences in the associations that we have found.

CONCLUSIONS

This large, long-term retrospective study found that both higher documentation heterogeneity and higher documentation intensity of lifestyle counseling are associated with faster achievement of A1c control. The findings suggest that these two quantitative characteristics of lifestyle counseling documentation may reflect the quality, quantity, and/or effectiveness of lifestyle counseling provided. Heterogeneity and intensity of documentation of lifestyle counseling in narrative provider notes may be used as indicators of quality of diabetes care.
REFERENCES


