Exploration of Temporal ICD Coding Bias Related to Acute Diabetic Conditions

Mollie McKillop, M.P.H., Fernanda Polubriaginof, M.D., Chunhua Weng, Ph.D.
Department of Biomedical Informatics, Columbia University, New York, NY, USA

Abstract
Electronic Health Records (EHRs) hold great promise for secondary data reuse but have been reported to contain severe biases. The temporal characteristics of coding biases remain unclear. This study used a survival analysis approach to reveal temporal bias trends for coding acute diabetic conditions among 268 diabetes patients. For glucose-controlled ketoacidosis patients we found it took an average of 7.5 months for the incorrect code to be removed, while for glucose-controlled hypoglycemic patients it took an average of 9 months. We also examined blood glucose lab values and performed a case review to confirm the validity of our findings. We discuss the implications of our findings and propose future work.

1.0 Introduction
Administrative data, namely ICD (International Classification of Diseases) codes, have been widely used to identify disease specific cohorts of patients. Such codes are often used in clinical and health services research because they are easy to obtain, have low associated costs, and can be aggregated to form large study samples. The accuracy of study results derived from administrative data depends on how well a particular coding scheme can correctly describe a disease cohort of interest. It is especially important to document coding biases as EHRs become more widely adopted and relied upon for large-scale data reuse. Increasingly, EHRs are used to document macroscopic human conditions, or phenotypes, automatically feeding data for secondary re-use purposes such as clinical research, quality improvement, and public health initiatives. Such uses require high-quality data, which are often lacking in the EHR. Coding bias is important to document and characterize for diabetes, which is an increasingly prevalent disease and a major source of morbidity and mortality. It is a leading cause of blindness, end-stage renal disease and cardiovascular disease and is associated with high healthcare costs. Coding bias related to complications of diabetes, specifically ketoacidosis and hypoglycemia, are particularly important to understand because they are acute, life-threatening conditions that require hospitalization. Accurate results from studies using the EHR to phenotype these patients must be aware of any coding bias related to these conditions.

The validity of ICD-codes for identifying patient groups has been challenged many times before and for a variety of conditions. These studies have shown that ICD codes are biased because concept definitions for codes are incomplete or are unsatisfactory in granularity. Moreover, variability in coding behavior also leads to incorrect code assignment. Researchers have previously questioned the validity and generalizability of ICD codes as applied to diabetes. However, these studies have not examined coding bias among complications of diabetes and uncontrolled glucose levels. To date, we could only find two studies that examined coding bias associated with ketoacidosis. They are small in scale and focus on data captured before 2009. We seek to build upon this work by validating these results and examining bias among a larger, more diverse, more recently treated population of patients. We also seek to report coding bias related to hypoglycemia, which to our knowledge has not previously been documented. Recognizing that human phenotypes are time-dependent, we aim to describe the temporal bias associated with ketoacidosis and hypoglycemia ICD-9 codes. Previous research has neglected the dynamic nature of such phenotypes when examining coding bias. Temporal bias is important to understand for accurate phenotyping and the full-realization of purported EHR benefits.

This study uses acute complications of diabetes and uncontrolled glucose as examples to investigate the temporality of ICD coding bias. Specifically, we ask the question, “do patients, who are initially coded for ketoacidosis or hypoglycemia, remain coded as such, despite controlling their glucose levels?” In other words, we hypothesize that patients who initially receive an ICD-9 code assignment for either ketoacidosis or hypoglycemia continue to be assigned these codes, despite little clinical evidence that such code assignment is reasonable. We examine the extent of this bias over time by using survival analysis to determine the time it takes, on average, for patients to have the incorrect code removed from their personal EHR. We also examine this bias for different disease subgroups (Type 1 versus Type 2) and among glucose-controlled and uncontrolled patients. Finally, we hope to describe our method in enough detail to allow other researchers to replicate our findings and test the strength of our conclusions by examining temporal coding bias among other acute conditions. This study was performed in compliance with the
World Medical Association Declaration of Helsinki on Ethical Principles for Medical Research Involving Human Subjects and was approved by the Columbia University Medical Center Institutional Review Board.

2.0 Methods

2.1 Data description and processing

For our study we utilized the Columbia University Medical Center Clinical Data Warehouse (CDW), which contains 24 years of data on 4.5 million patients\textsuperscript{11}. Data were extracted in a multi-step process as depicted in Figure 1 to generate our cohort for analysis. First, relevant ICD-9 codes were identified using the Center for Medicare and Medicaid (CMS) ICD-9 Code Lookup\textsuperscript{12}. The relevant codes are listed in Table 1. All instances of these codes from 2004 to October 2014 along with corresponding fake patient-identification numbers and real timestamps for each code were extracted. This resulted in 4844 unique patients, 2242 with ketoacidosis and 2602 with hypoglycemia. Along with the ICD-9 codes, all HgA1c lab values for these patients along with fake patient IDs and real timestamps were generated from the CDW. Lab values occurring after the initial coding of ketoacidosis or hypoglycemia were selected. Patients were then separated by ketoacidosis and hypoglycemia. After selecting only the lab values occurring after the initial coding for ketoacidosis, 1309 patients were retained; after selecting only the lab values occurring after the initial coding for hypoglycemia, 1129 patients remained. In the ketoacidosis subgroup, 112 patients were selected for analysis based on having controlled glucose levels, which was defined as having a median HgA1c lab value less than 7, a threshold recommended by the American Diabetes Association\textsuperscript{13}. In the hypoglycemia group, 156 patients were selected for analysis based on having a similarly defined controlled glucose level.

2.2 Data analysis

Survival analysis was performed separately on the remaining 112 ketoacidosis and 156 hypoglycemia observations to determine the probability of remaining coded for ketoacidosis or hypoglycemia, despite having controlled glucose levels. Survival analysis was chosen as the statistical method for analysis because it allows for the time to an event to be computed. In this case, the event was decoding of a glucose-controlled patient. For the ketoacidosis group, the granularity of the ICD-9 codes allowed us to further explore any coding bias among disease subgroups.

First, individuals who were identified as Type 1 or Type 2 ketoacidosis patients by ICD-9 codes had their time to decoding compared and assessed for statistical significance using a cox proportional hazards test. Second, individuals who were identified as uncontrolled or controlled ketoacidosis patients by ICD-9 codes had their time to decoding compared and assessed for statistical significance again using a cox proportional hazard test. Because HgA1c lab tests measure a patient’s average glucose level over several weeks, we looked at the blood glucose lab tests 14 days before any ICD-9 code timestamp after the first ICD-9 code assignment. This time frame was chosen based on administrative timetables, since reimbursement claims are generally submitted within two weeks of any services or procedures provided. Also, the ICD-9 timestamp represents the discharge date. Considering the acute nature of the diseases under study, we felt this was an adequate time frame to select blood glucose lab results from.

The blood glucose test was selected because it provides a way to characterize patients as either in a state of hypoglycemia or ketoacidosis. According to UpToDate, a blood glucose serum level over 350 can be used to diagnose ketoacidosis\textsuperscript{14} while a blood glucose serum level under 40 can be used to diagnose hypoglycemia\textsuperscript{15}. The
blood glucose lab values were extracted from the CDW using the Medical Entities Dictionary, which is a large repository of medical concepts that are drawn from a variety of sources either developed or used at NYP. For each patient, the mean and median percentage of times these blood glucose lab values indicated disease for either the ketoacidosis or the hypoglycemia group was calculated. The analysis was performed using Rstudio version 0.98.1091. The datasets and R code are available for public use by contacting the primary author (MM).

Finally, a case review of deceased patient notes for both the ketoacidosis group and the hypoglycemia group was performed by a physician (FP) to further examine the validity of our results and perform an error analysis. Since patients were selected from groups dependent upon ICD-9 code assignment, the clinical reviewer provided a comprehensive chart review of the deceased patient notes from both the ketoacidosis and hypoglycemia group. The main purpose of this review was to contribute to an error analysis and provide a definitive “gold standard” of whether the patient had ketoacidosis or hypoglycemia. The Annals of Emergency Medicine guidelines for chart review were followed as much as possible. The participating reviewer (FP) is a clinician and knew the purpose of the study but not the study’s full outcome. The variables to be collected from the chart, as well as how these variables are defined, were determined a priori and documented for the coder. Variables were selected using UptoDate Guidelines to construct coding rules. Once a positive diagnosis based on the coded variables was determined by the reviewer, no further case review for that patient was performed. Overall, the reviewer examined 5 ketoacidosis patients and 13 hypoglycemic patients.

3.0 Results
3.1 Descriptive statistics
In total there were 112 patients in the ketoacidosis group and 156 patients in the hypoglycemia group. All patients that occurred in the ketoacidosis group also occurred in the hypoglycemia group. In the ketoacidosis group, there were a median of 5 HgA1c lab values per patient with a minimum of 1 and a maximum of 40. In the hypoglycemia group, there were a median of 10 HgA1c lab values per patient with a minimum of 2 and a maximum of 203. The trend in HgA1c lab values over time by patient for the ketoacidosis group (left) and hypoglycemia group (right) is shown in Figure 2.

![Ketoacidosis Patients HgA1C Linear Trend](image1)

![Hypoglycemia Patients HgA1c Linear Trend](image2)

Figure 2. HgA1c linear trend for ketoacidosis patients (left) and hypoglycemia patients (right). Each line represents a unique patient and colors are used to help distinguish different lines. Horizontal axis is sequential order of HgA1c lab tests; Vertical axis is HgA1c lab value.

A generalized estimating equation, assuming data were missing completely at random, was fit for the ketoacidosis and hypoglycemia groups separately to test for significant linear trend associations between HgA1c and time. The hypoglycemia group was found to have a significant association (p-value = 0.0045) while the ketoacidosis did not have such an association (p-value = 0.11). Both tests had a negative coefficient. In the ketoacidosis group, there were a median of 15.5 ICD-9 code assignments per patient with a minimum of 1 and a maximum of 211. In the hypoglycemia group, there were a median of 10 ICD-9 code assignments per patient with a minimum of 2 and a maximum of 203. Among the ketoacidosis group split by diabetes type as determined by ICD-9 codes, there were 39 Type 1 patients and 73 Type 2 patients. Among the ketoacidosis group split by uncontrolled versus controlled diabetes type as determined by ICD-9 codes and not median HgA1c lab values, there were 53 ‘not stated as uncontrolled’ patients and 59 ‘uncontrolled’ patients.

3.2 Survival analysis for the ketoacidosis group
The Kaplan-Meier survival curve for the 112 observations is shown in Figure 3. Time was recorded in 3-month intervals to reflect national guidelines for diabetes screening. According to the survival curve, it takes approximately 7.5 months for 50% of patients, who were initially coded for ketoacidosis and have their glucose in control, to stop being assigned any of the four ICD-9 codes listed in Table 1.

**Figure 3.** Kaplan-Meier survival curve for all ketoacidosis patients. Time is in 3-month intervals. Dashed lines represent confidence intervals for the survival curve.

A cox-proportional hazard test was performed to determine if significant differences existed between any of the four ICD-9 codes. This test was found to be significant at a level of .01 (p-value=0.00538), indicating at least one of the ICD-9 code survival curves is different from the others. This difference was examined further by separating the population into different groups based on diabetes type and ‘uncontrolled’ versus ‘not stated as uncontrolled’ ICD-9 code assignment. The survival curves for Type 1 and Type 2 diabetes are displayed in Figure 4. These curves were not significantly different from each other at a level of .05 (p-value=0.316) using the cox-proportional hazards test.

**Figure 4.** Kaplan-Meier survival curve for ketoacidosis patients by diabetes type. Time is in 3-month intervals. The blue curve represents Type 1 diabetics and green curve represents Type 2 diabetics.

A third cox-proportional hazard test determined a significant difference (p-value=.00821) at a level of .01 between the Kaplan-Meier survival curves of patients coded for ‘uncontrolled’ (ICD-9 codes 250.12 and 250.13) versus
patients coded for ‘not stated as uncontrolled’ (ICD-9 codes 250.10 and 250.11). The survival curves are displayed in Figure 5 and indicate patients coded for ‘not stated as uncontrolled ketoacidosis’ (blue line) are decoded from any ketoacidosis ICD-9 code 25.1% faster than patients coded for ‘uncontrolled ketoacidosis’ (green line).

![Survival Curve of 'Not Stated as Uncontrolled' vs. 'Uncontrolled']

**Figure 5.** Kaplan-Meier survival curve for ketoacidosis patients by ICD-9 code assignment for ‘Not Stated as Uncontrolled’ (ICD-9 codes 250.10 and 250.11) vs. ‘Uncontrolled’ (ICD-9 codes 250.12 and 250.13). Time is in 3-month intervals. The blue curve is for ‘Not Stated as Uncontrolled’ diabetics and the green curve is for ‘Uncontrolled’ diabetics.

### 3.3 Glucose lab value assessment for the ketoacidosis group

Patients were selected from those that had one or more ICD-9 code assignments for ketoacidosis after the initial coding for the disease. The glucose lab tests occurring 14 days before any ICD-9 code timestamp, except the first timestamp, were selected and assessed for positive indication of ketoacidosis based on a blood glucose cutoff level of 350 mg/dL as determined by UpToDate\(^4\). Only 90 out of the 112 patients had blood glucose labs occurring 14 days before the ICD code timestamp values under consideration. Positive disease indications were assigned a value of one and tabulated by patient. The mean number of positive disease indications was 210 with a median of 195, while the median percentage of positive disease indications was 6.33% of all blood glucose lab tests per patient. The mean percentage of positive disease indications over the number of blood glucose labs was calculated per patient and the distribution is shown in Figure 6.

![Distribution of mean positive ketoacidosis indication by sequential ICD code assignments.](image)

**Figure 6.** Distribution of mean positive ketoacidosis indication by sequential ICD code assignments.

2009
3.4 Survival analysis for the hypoglycemia group
The Kaplan-Meier survival curve for the 156 hypoglycemia patients is shown in Figure 7. Time was recorded in 3-month intervals to reflect national guidelines for diabetes screening. According to the survival curve, it takes 9 months for 50% of patients, who were initially coded for hypoglycemia and have their glucose in control, to stop being assigned any of the ICD-9 codes for ketoacidosis. Unfortunately, no further survival analysis of hypoglycemia subgroups could be performed due to the limited granularity of ICD-9 codes in this category.

![Kaplan-Meier survival curve for all hypoglycemia patients. Time is in 3-month intervals. Dashed lines represent confidence intervals for the survival curve.](image)

**Figure 7.** Kaplan-Meier survival curve for all hypoglycemia patients. Time is in 3-month intervals. Dashed lines represent confidence intervals for the survival curve.

3.5 Glucose lab value analysis for the hypoglycemic group
Patients were selected from those that had one or more ICD-9 code assignments for hypoglycemia after the initial coding for the disease. The glucose lab tests occurring 14 days before any ICD-9 code timestamp, except this first timestamp, were selected and assessed for positive indication of hypoglycemia based on a blood glucose cutoff level of 40 mg/dL as determined by UpToDate\(^5\). Only 33 out of the 156 patients had blood glucose labs occurring 14 days before the ICD code timestamp values under consideration. Positive indications were assigned a value of one and tabulated by patient. The mean number of positive disease indications was 120 with a median of 109, while the mean percentage was .8977% and the median percentage was 0%. The percentage of positive disease indications over the number of blood glucose labs was calculated per patient and the distribution is shown in Figure 8.
4.0 Discussion

4.1 Major Implications

The results of this study seem to suggest a coding bias related to ICD-9 ketoacidosis and hypoglycemia codes. It appears that those who are initially coded for ketoacidosis or hypoglycemia continue to be coded as such despite having their glucose levels in control, as indicated by median HgA1c lab values. This finding is also supported by the low median percentage of blood glucose lab tests indicating disease for the ketoacidosis group (6.33%) and the hypoglycemia group (0%). More importantly, this study examines the temporal nature of such coding bias. Because ketoacidosis and hypoglycemia are conditions that last several hours to days, our results show the extent of this bias seems severe when the temporality of coding behavior is considered. This bias persists when diabetes type and control, as determined by ICD-9 codes, are examined. For example, patients that are ICD-9 coded for ‘uncontrolled diabetes, ketoacidosis’ and have controlled glucose levels (as determined by HgA1c lab tests) have their inaccurate code removed 79.9% slower than the same patients coded for ‘not stated as uncontrolled diabetes, ketoacidosis’. This makes clinical sense considering those that have uncontrolled diabetes have a higher probability for ketoacidosis, a complication related to poor glucose control. With regard to hypoglycemia, temporal coding bias seems to remain and the effect is even stronger than in the ketoacidosis group. This may be due to differing disease severity in that hypoglycemia is a less severe disease than ketoacidosis, so more people are likely to be coded for its presence. However, it remains unclear why patients, who have lab values indicating their glucose is already controlled, continued to receive hypoglycemia codes. We hypothesize that because hypoglycemia is a less severe and a less formally defined disease than ketoacidosis, it is easier to interpret this diagnosis from patient notes and lab values. This allows for coders to more frequently assign hypoglycemia ICD-9 codes. Our findings are similar to previous studies that found ICD-9 coding bias among ketoacidosis patients. Moreover our results imply a potential overestimation of disease prevalence if ICD-9 codes are used solely to identify a ketoacidosis or hypoglycemia cohort.

4.2 Error analysis

In the ketoacidosis group all 5 patients were positively identified by manual case review as having ketoacidosis. In the hypoglycemia group, 5 patients did not have notes at any point two weeks before the first ICD code assignment, leaving 8 code assignments for our error analysis. It remains unknown as to why these notes are missing. Of these 8 assignments, 7 were confirmed as correct (patients had hypoglycemia) with the remaining incorrect code due to a patient having normal glucose levels. Major sources of error therefore include patients with normal glucose being coded for hypoglycemia as well as no clinical basis for code assignment, since no notes existed for several patients.

In addition to the error analysis, we also identified deceased patients who had repeated ICD-9 code assignments for their particular disease group. The purpose of this assessment was to see if repeated ICD-9 code assignments were
valid. Only one patient fit this description in the ketoacidosis group. In a case review performed in the same manner as the error analysis, the patient was identified as having normal glucose. In the hypoglycemia group, all 13 patients had repeated ICD-9 code assignments, with a total of 16 repeated assignments (one patient had 4 repeated assignments after the first assignment, while the rest of the patients had one). This analysis revealed that 8 of these ICD code assignments were missing all notes two weeks before the code assignment. Of the 8 remaining code assignments, 6 of these were found to be correct through case review of patients’ notes. The other two assignments for hypoglycemia were assigned to patients with normal glucose.

The survival analysis and mean blood glucose results along with the case review suggest a semantic gap between the physicians who author such notes and the coders who ultimately assign ICD-9 codes. The physician’s concept of the patient must be encoded in the patient notes and then from this computable representation the coder must form a concept of an ICD-9 code. In this process it appears that clinical information is lost or misunderstood. Either the concept is not encoded (i.e no notes are present two weeks before a code assignment) or the concept is semantically altered once present in the EHR. The coding bias we describe could be partially explained by the over-use of copy-pasted text, which appeared multiple times in each of the case-reviewed patient notes. Finally, economic incentives for incorrectly coding patients should also be considered and further studied as a potential source of coding bias.

4.3 Limitations

This study has several limitations. First, the sample size of 112 ketoacidosis patients and 156 hypoglycemia patients is relatively small, and larger studies could reveal interesting differences in ketoacidosis coding between Type 1 and Type 2 patients. Moreover a larger dataset could allow for more complex measures of glucose control. Because the median number of HgA1c lab tests performed was only 5 for the ketoacidosis group, more sophisticated approaches, such as time series analysis, that can measure trends in glycemic control were unfeasible. Moreover, our goal was not to model HgA1c trends but to characterize a group of patients with reasonably well-controlled glucose levels. Therefore, median HgA1c lab values were used as a measure for glucose management. However, we feel this is a reasonable assessment of glucose control in relation to the acute complications we are exploring. Finally, it should be noted that all HgA1c measurements were collected observationally, and that there may be high variability in measurement quality.

In addition, it is possible our findings are due to repeated hospitalizations for patients who do in fact have either hypoglycemia or ketoacidosis, yet this seems unlikely given the low median percentage of glucose blood tests indicating either disease and the follow up we performed in our error analysis. Finally, the two groups of patients are not independent from each other, in that all ketoacidosis patients are also represented in the hypoglycemia group. This limits the generalizability of our findings to other patient groups.

Importantly, there may be outliers influencing the results of our analysis, as is implied by the relatively large difference between the number of maximum ICD-9 code assignments and median number of ICD-9 code assignments for both the ketoacidosis group and the hypoglycemia group. We did do a visual evaluation to identify these points, plotting the patient id against the number of ICD code assignments for both disease groups. There appeared to be two outliers among the ketoacidosis group and another two outliers in the hypoglycemia group. We did not remove these outliers since our aim was exploratory in nature and there appeared to be few outliers. In addition, we do not wish to propose a mathematical model for temporal coding bias but rather to prove the feasibility of a method for exploring this phenomenon. More robust and larger datasets could aid in determining why such outliers might exist.

Only 18 deceased patients’ records (5 in the ketoacidosis group and 13 in the hypoglycemia group) were reviewed. These patients may not represent each disease group accurately. Because only deceased patients’ notes were used for the error analysis, bias may be introduced by selecting patients who are perhaps more likely to die, and therefore more sick, than other patients. However, we feel this bias is limited, since its presence would shift our results toward concluding that there is less coding bias (as it is more likely that coders would assign ICD-9 ketoacidosis and hypoglycemia diagnosis codes). On the other hand, these patients may be so sick that coders fail to identify ketoacidosis and hypoglycemia as the patient’s primary reason for hospital visit in favor of codes representing more serious diagnoses. Yet, ketoacidosis and hypoglycemia are very serious life-threatening conditions that would probably outrank most other diseases in terms of prominence and importance during a hospital visit.

Finally, the numeric results of this study are probably not generalizable to other settings, because decoding of ketoacidosis could be due to either the coder stopping ketoacidosis code assignment or, perhaps more likely, the patient not returning to the New York-Presbyterian Hospital for care. In either case, the extent of this bias would alter our results in the direction of less temporal coding bias.
4.4 Future work

Future work should try to examine the reasons for the documentation discrepancies between physicians and coders as well as any financial motivations coders might have for assigning incorrect codes. Physicians could try to be more consistent in their terminology and use less copy-pasted text, so that coders have a clearer representation of the ketoacidosis and hypoglycemia concept. Coders and physicians could be better educated on the importance of communicating and picking accurate ICD-9 codes. In addition, more research should be done to develop accurate and automated methods of ICD-9 code assignment, which could overcome the coding discrepancies between physician and coder. This study also reveals challenges in secondary reuse of clinical data for research purposes. More relevant organization of clinical data would allow for easier data analysis and increase research efficiency. For example, this could occur through front-end applications that allow users to view, sort and analyze clinical data.

Conclusion

The results of this study indicate that temporal coding bias is a problem among patients with acute complications related to poor glucose control. This bias is significantly different between those coded as 'not stated as uncontrolled' vs. 'uncontrolled' with ICD-9 ketoacidosis codes and occurs among two different acute complications, i.e., ketoacidosis and hypoglycemia. A case review of patients’ notes further confirmed this bias. We also contribute a novel method for coding bias research related to acute diseases. Understating such bias is increasingly important as the EHR becomes more widely used for phenotyping and cohort selection. Future work should have a socio-technical approach, evaluating how these biases arise. Specifically, the coding discrepancies between physicians and coders should be further explored and documented.

Acknowledgements

This study was sponsored by the U.S. National Library of Medicine grants R01LM009886 (PI: Weng) and T15LM007079 (PI: Hripcsak) and U.S. National Center for Advancing Translational Science grant UL1 TR000040 (PI: Ginsberg).

References


