Received: 21 September 2021 | Accepted: 16 February 2022

DOI: 10.1111/dme.14815

DIABETIC Medicine

RESEARCH: COMPLICATIONS

Baseline haemoglobin A1c and the risk of COVID-19 hospitalization among patients with diabetes in the INSIGHT Clinical Research Network

Jea Young Min[1] | **Nicholas Williams**[1] | **Will Simmons**[1] | **Samprit Banerjee**[1] | **Fei Wang**[1] | **Yongkang Zhang**[1] | **April B. Reese**[3] | **Alvin I. Mushlin**[1] | **James H. Flory**[1,2]

[1]Department of Population Health Sciences, Weill Cornell Medical College, New York, New York, USA

[2]Memorial Sloan Kettering Cancer Center, New York, New York, USA

[3]Patient Stakeholder Partner, Independent Contractor

Correspondence
James H. Flory, 530 East 74th Street, New York, NY 10021, USA.
Email: floryj@mskcc.org

Funding information
National Cancer Institute (NCI); Patient-Centered Outcomes Research Institute

Abstract

Aims: To examine the association between baseline glucose control and risk of COVID-19 hospitalization and in-hospital death among patients with diabetes.

Methods: We performed a retrospective cohort study of adult patients in the INSIGHT Clinical Research Network with a diabetes diagnosis and haemoglobin A1c (HbA1c) measurement in the year prior to an index date of March 15, 2020. Patients were divided into four exposure groups based on their most recent HbA1c measurement (in mmol/mol): 39–46 (5.7%–6.4%), 48–57 (6.5%–7.4%), 58–85 (7.5%–9.9%), and ≥86 (10%). Time to COVID-19 hospitalization was compared in the four groups in a propensity score-weighted Cox proportional hazards model adjusting for potential confounders. Patients were followed until June 15, 2020. In-hospital death was examined as a secondary outcome.

Results: Of 168,803 patients who met inclusion criteria; 50,016 patients had baseline HbA1c 39–46 (5.7%–6.4%); 54,729 had HbA1c 48–57 (6.5–7.4%); 47,640 had HbA1c 58–85 (7.5^%–9.9%) and 16,418 had HbA1c ≥86 (10%). Compared with patients with HbA1c 48–57 (6.5%–7.4%), the risk of hospitalization was incrementally greater for those with HbA1c 58–85 (7.5%–9.9%) (adjusted hazard ratio [aHR] 1.19, 95% confidence interval [CI] 1.06–1.34) and HbA1c ≥86 (10%) (aHR 1.40, 95% CI 1.19–1.64). The risk of COVID-19 in-hospital death was increased only in patients with HbA1c 58–85 (7.5%–9.9%) (aHR 1.29, 95% CI 1.06, 1.61).

Conclusions: Diabetes patients with high baseline HbA1c had a greater risk of COVID-19 hospitalization, although association between HbA1c and in-hospital death was less consistent. Preventive efforts for COVID-19 should be focused on diabetes patients with poor glucose control.

KEYWORDS

COVID-19, diabetes mellitus, glycated haemoglobin A, hospitalization, mortality

1 | INTRODUCTION

Recent studies suggest that diabetes is a risk factor for COVID-19 and poor outcomes including hospitalization and death.[1-6] Patients with diabetes are generally at a greater risk of serious infections due to mechanisms including altered immune response, altered metabolism and diabetic complications.[7] Some studies also suggest that hyperglycaemia can lead to an elevated expression of angiotensin-converting enzyme 2 (ACE2), which could contribute to worse SARS-CoV-2 infection given the pathogen's affinity for that receptor.[8,9] However, diabetes is also associated with other comorbidities and risk factors for poor outcomes, potentially confounding any observed association with COVID-19 outcomes.

A key question concerning the interplay between diabetes and COVID-19 is whether the risk of severe COVID-19 among diabetic patients differs depending on the level of glucose control at the time of exposure. As the level of haemoglobin A1c (HbA1c) varies widely among patients with diabetes, this question can be addressed by leveraging large secondary databases, with appropriate methods to control for potential confounders. Previously, a population-based cohort study in the United Kingdom found that the risk of COVID-19 mortality was elevated among patients with higher levels of HbA1c prior to the COVID-19 pandemic, but the study included relatively limited adjustment for potential confounders as it explored multiple risk factors in the same model.[10] Other small studies that used data from single hospitals and only included patients already sick enough to be hospitalized, found conflicting results regarding the association between HbA1c or blood glucose levels on admission and the severity of COVID-19.[11-15] The result is an important evidence gap on a fundamental question: does poor outpatient glycaemic control lead to an increased risk of severe COVID-19 disease?

The objective of our study was to examine the association between baseline HbA1c and the risk of COVID-19 hospitalization in a large cohort of patients with diabetes.

Novelty Statement

- It is unclear whether the risk of severe COVID-19 outcomes varies depending on the level of baseline glucose control among patients with diabetes.
- Diabetes patients with high baseline HbA1c had a greater risk of COVID-19 hospitalization compared with those with optimal glucose control although associations between HbA1c and in-hospital death were less consistent.
- Preventive efforts for COVID-19 should be focused on diabetes patients with poor glucose control.

2 | METHODS

2.1 | Study design and data sources

We conducted a retrospective cohort study using data from the INSIGHT Clinical Research Network (CRN), formally known as the New York City Clinical Data Research Network (NYC-CDRN).[16] The INSIGHT CRN data are comprised of electronic health record (EHR) data from an 11-year period that includes longitudinal data for a large, diverse urban patient population across five academic medical centres in New York City: Albert Einstein School of Medicine/Montefiore Medical Centre, Columbia University and Weill Cornell Medicine/New York-Presbyterian Hospital, Icahn School of Medicine/Mount Sinai Health System, Clinical Director's Network, and New York University School of Medicine/Langone Medical Centre.[17] The CRN data include information on demographics, social determinants of health based on residential zip code, inpatient and outpatient health care encounters, clinical measurements, laboratory measurements and electronic prescriptions for patients with records at one or more of the participating health systems.

This study was approved by the Weill Cornell Medicine and Memorial Sloan Kettering Cancer Center Institutional Review Boards and the requirement to obtain informed consent from the study participants was waived.

2.2 | Study population

We initially obtained data from all time periods for patients in the INSIGHT CRN with potential evidence of diabetes mellitus, including those with a diagnosis code for diabetes, a prescription for an antidiabetic drug, or a HbA1c >48 mmol/mol (6.5%). Of this initially selected cohort, patients aged 18 years or older who were alive on the index date of March 15, 2020 and had an HbA1c measurement in the year prior to the index date were included. The index date was chosen as the date on which exponential increases in COVID-19 cases in New York City were observed.[18] We excluded patients with a HbA1c measurement <39 mmol/mol (5.7%) and those without a diagnosis of type 1 or type 2 diabetes (International Classification of Diseases, 9th Revision, Clinical Modification [ICD-9-CM] 250.X or ICD-10-CM E10.X-E14.X). These requirements

were intended to select a population who were likely receiving regular care within the healthcare system and also to ensure that the population did not include patients with pre-diabetes or obesity alone.

2.3 | Baseline HbA1c groups

Study patients were classified into four groups based on their baseline HbA1c (in mmol/mol), defined as their most recent HbA1c measurement in the year prior to the index date: 39–46 (5.7%–6.4%), 48–57 (6.5%–7.4%), 58–85 (7.5%–9.9%) and ≥86 (10%). Patients with baseline HbA1c 48–57 mmol/mol (6.5%–7.4%) were treated as the reference group. Patients were followed until June 15, 2020 and were censored if they died prior to the end of follow-up.

2.4 | Study outcomes

The primary outcome of the study was COVID-19 hospitalization, defined by a discharge diagnosis code for COVID-19 and/or a positive COVID-19 test result either during or 2 weeks prior to the hospitalization (Table S1). In-hospital death was examined as a secondary outcome.

We also examined characteristics of COVID-19 hospitalizations that could be indicators of disease severity and pathogenesis, including comorbid ketoacidosis, acute kidney injury (AKI), and hypoglycaemia.[19,20] Diagnoses were identified using ICD-10-CM discharge diagnosis codes from a COVID-19 hospitalization (Table S1). For AKI and hypoglycaemia, laboratory-based definitions were also used to identify possible occurrences. Laboratory-based AKI was defined if patients had a highest creatinine measurement during the hospitalization that was >44.2 μmol/L (0.5 mg/dl) greater than their baseline, >2 times greater their baseline creatinine level, or >353.6 μmol/L (4 mg/dl).[21] Laboratory-based hypoglycaemia was defined if there was a blood glucose measurement <2.78 mmol/L (50 mg/dl) during the hospitalization, which has been accepted as clinically significant sign of hypoglycaemia regardless of symptoms in diabetes guidelines and previous clinical trials.[22-24]

2.5 | Covariates

We selected 36 covariates a priori as potential confounders in the association between HbA1c and COVID-19 hospitalization. Covariate information was collected from up to 1 year prior to the most recent HbA1c measurement before the index date and included demographics (age, sex, race, ethnicity, social deprivation index[25]), and baseline vital signs and laboratory measurements (body mass index [BMI], systolic blood pressure, diastolic blood pressure, serum creatinine]. Comorbidities were identified using the Elixhauser algorithm.[26-28] We included the ten most common comorbidities that were present in the study population: uncomplicated and complicated hypertension, obesity, renal failure, chronic pulmonary disease, cardiac arrhythmia, hypothyroidism, fluid and electrolyte disorders, congestive heart failure and depression (Table S2). Medication use during the baseline year was defined as having a prescription for any of the following drug classes: antidiabetic drugs (insulin, metformin, sulfonylureas, dipeptidyl peptidase-4 [DPP-4] inhibitors, sodium glucose co-transporter 2 [SGLT2] inhibitors, glucagon-like peptide 1 [GLP-1] agonists, thiazolidinediones), antihypertensive drugs, aspirin, statins, immunosuppressants, antidepressants and antipsychotics (Table S2). Healthcare utilization metrics included the number of inpatient admissions, outpatient encounters, emergency department encounters and outpatient medications during the baseline year.

2.6 | Statistical analysis

We compared time to COVID-19 hospitalization in four HbA1c groups in a Cox proportional hazards model with inverse-probability of treatment weighting (IPTW) using estimated propensity scores and additional covariate adjustment. Propensity scores estimated the probability of being in each HbA1c category based on covariates and were estimated using an ordinal logistic regression model with HbA1c 48–57 mmol/mol (6.5%–7.4%) as the reference group (Table S3; Figure S1).[29] Multiple imputation was conducted (30 imputations) for missing covariates using Multivariate Imputation by Chained Equations (MICE) method.[30] Continuous variables were modelled using restricted cubic splines to account for nonlinearity. The secondary outcome, in-hospital death, was also compared in the four HbA1c groups in a similar Cox proportional hazards model with IPTW and covariate adjustment. All analyses were conducted using R 3.6.2.

2.7 | Sensitivity and subgroup analyses

We performed sensitivity analyses to check the robustness of our findings. First, we used a data-adaptive machine learning approach, specifically the SuperLearner algorithm,[31] to develop the propensity score weights. Data-adaptive algorithms, such as the SuperLearner, may converge to the true answer at a rate slower than $n^{-1/2}$, leading to under-coverage of confidence

intervals; confidence intervals were, therefore, not estimated.[32] The following algorithms were included in the SuperLearner: an intercept-only model, main-effects GLM, Multivariate Adaptive Regression Splines (MARS), L1 penalization (LASSO) and eXtreme Gradient Boosting.[31]

Second, we limited the study sample to those with a high predicted EHR continuity score, or mean proportion of encounters captured by the EHR system. One of the limitations of using CRN data is that we may miss many hospitalization outcomes that occur outside the CRN network. We applied a modified version of the EHR data continuity prediction model developed by Lin et al.[33] to linked INSIGHT-CRN and Medicare data from 2013–2016, excluding variables not reliably available in the INSIGHT-CRN and confirming through cross-validation that it was able to distinguish high data-continuity from low data-continuity patients (Table S4). We then used the modified algorithm to limit our population to those with a predicted EHR continuity score of 60% or greater.

We also performed subgroup analyses to assess the association between HbA1c and risk of COVID-19 hospitalization after stratification by sex, race/ethnicity and age (<65 and ≥65 years).

2.8 | Exploratory analysis

To explore whether the association between HbA1c and the primary outcome may be attributable to a unique COVID-19-related mechanism, we examined influenza hospitalizations as a 'negative control' outcome in the same study population. The influenza hospitalization outcome was defined by a positive influenza A or B laboratory test result during or 2 weeks prior to a hospitalization (Table S1).

3 | RESULTS

3.1 | Cohort identification and patient characteristics

Of all 1,059,301 patients identified in the CRN cohort with potential evidence of diabetes during any time period, 241,008 patients had at least one HbA1c measurement between March 15, 2019, and March 14, 2020 (Figure 1). Of these patients, 775 were excluded because they did not meet the age requirement, and 2,157 died before the index date. An additional 29,522 patients were excluded because they did not have a HbA1c measurement of 39 mmol/mol (5.7%) or higher, and 39,751 were

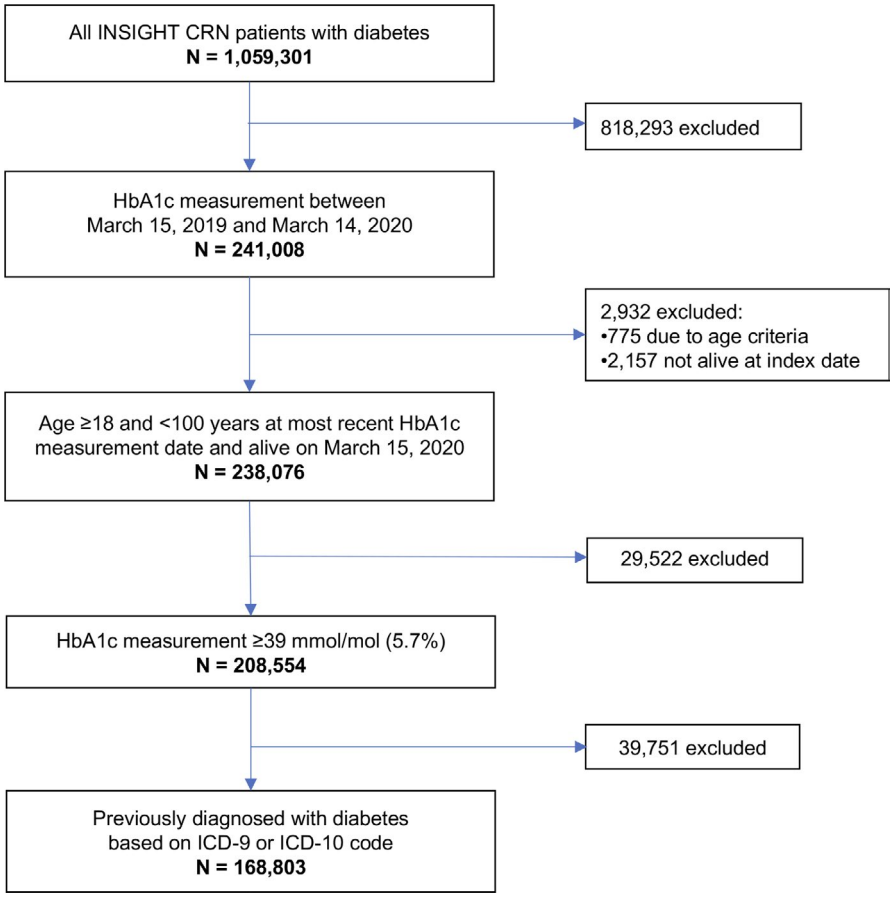

FIGURE 1 INSIGHT Clinical Research Network (CRN) patients eligible for study inclusion; the initial cohort of CRN patients with diabetes includes patients in the INSIGHT CRN with potential evidence of diabetes from any time period

excluded because they did not have a previous diagnosis of diabetes in the CRN data.

Of 168,803 patients who met all the inclusion criteria, there were 50,016 patients with a baseline HbA1c (in mmol/mol) between 39 and 46 (5.7%–6.4%), 54,729 patients with HbA1c 48–57 (6.5%–7.4%), 47,640 with HbA1c 58–85 (7.5%–9.9%) and 16,418 with HbA1c ≥86 (10%). Patients in the lower HbA1c groups were more likely to be female, white, non-Hispanic, and older compared with those in the high HbA1c group (Table 1). Propensity score weighting using propensity scores derived from logistic regression improved the balance of covariate distributions across the low HbA1c groups as shown by the standardized mean differences (Table S5). The sensitivity analysis that estimated propensity scores using the SuperLearner algorithm was superior in achieving covariate balance across all HbA1c comparison groups (Table S6), with standardized mean differences <0.1 for all covariates.

3.2 | COVID-19 hospitalization and in-hospital death

The unadjusted rates of COVID-19 hospitalization were 3.22, 3.04, 4.23 and 5.38 per 1000 person-days in each HbA1c group, from lowest to highest (Table 2). Compared with patients with HbA1c 48–57 mmol/mol (6.5%–7.4%), the risk of hospitalization was greater for HbA1c 58–85 (7.5%–9.9%) (adjusted hazard ratio [HR] 1.19, 95% confidence interval [CI] 1.06–1.34) and HbA1c ≥86 (10%) (adjusted HR 1.40, 95% CI 1.19–1.64) (Table 2). The risk of hospitalization was not significantly different for patients with HbA1c 39–46 (5.7%–6.4%) (adjusted HR 0.92, 95% CI 0.82, 1.04). The hazard ratios were similar in the sensitivity analysis that used SuperLearner to estimate propensity score weights (Table 2).

3.3 | Mortality and co-morbid conditions during COVID-19 hospitalization

The risk of COVID-19 in-hospital death was higher only in the HbA1c 58–85 mmol/mol (7.5%–9.9%) group compared with the HbA1c 48–57 (6.5%–7.4%) group (adjusted HR 1.29, 95% CI 1.05, 1.59). The effect estimates were consistent in the sensitivity analysis that used SuperLearner propensity score weights.

AKI diagnoses were present in 883 (39.7%), ketoacidosis diagnoses in 121 (5.4%), and hypoglycaemia diagnoses in 206 (9.3%) of COVID-19 hospitalizations (Table 3). Laboratory-based definitions identified 865 (38.9%) of hospitalizations with evidence of AKI and 133 (6.0%) hospitalizations with severe hypoglycaemia. Although evidence of AKI was common across all HbA1c groups, evidence of ketoacidosis and hypoglycaemia was much more common in the higher HbA1c groups. Specifically, a ketoacidosis diagnosis was present in only 0.9% of COVID-19 hospitalizations among patients with HbA1c 39–46 (5.7%–6.4%), but for patients with HbA1c 58–85 (7.5%–9.9%) and ≥86 (10%), it was present in 6.8% and 13.4% of COVID-19 hospitalizations, respectively.

3.4 | Sensitivity and subgroup analyses

The sensitivity analysis excluding patients with low predicted EHR continuity was generally consistent with the primary analysis, especially for the group with HbA1c ≥86 mmol/mol (10%). However, the association became attenuated in the HbA1c 58–85 (7.5%–9.9%) group and was no longer different from the reference group with HbA1c 48–57 (6.5%–7.4%). Overall, the absolute rates of COVID-19 hospitalization were much higher in the subgroup with high EHR continuity, ranging from 10.8 to 15.5 events per 1000 person-days.

In subgroup analyses, the association between HbA1c and risk of COVID-19 hospitalization did not differ materially among different groups of patients based on sex, race or age (Table S7).

At the recommendation of peer reviewers, a number of ad-hoc sensitivity analyses testing the robustness of the hazard ratio estimates for the primary outcome were also conducted. These included adjustment for all Elixhauser comorbidities (rather than the 10 included in the primary analysis), inclusion of renin-angiotensin antagonists as a distinct drug-exposure covariate and inclusion of patients with HbA1c < 5.7 in the cohort. Results were unaffected by these changes (meaning that all significant findings remained significant, and no hazard ratio estimate changed by more than 5%).

3.5 | Influenza outcome (negative control)

The rates of influenza hospitalizations were much lower compared with the rates of COVID-19 hospitalizations during the study period and did not differ among HbA1c groups after covariate adjustment (Table 2).

4 | DISCUSSION

4.1 | Study summary

Diabetes patients with elevated HbA1c had a greater risk of COVID-19 adverse outcomes compared with those

TABLE 1 Baseline characteristics of cohort patients by baseline HbA1c (mmol/mol [%])[a].

Characteristic	39–46 mmol/mol (5.7–6.5%) *N* = 50,016	smd	48–57 mmol/mol (6.5–7.4%) *N* = 54,729	58–85 mmol/mol (7.5–9.9%) *N* = 47,640	smd	≥86 mmol/mol (≥10%) *N* = 16,418	smd
Age	66 (13)	0.03	66 (13)	64 (14)	0.16	58 (15)	0.61
Sex							
Female	28174 (56%)	0.07	28822 (53%)	24249 (51%)	0.04	8099 (49%)	0.07
Male	21838 (44%)	0.07	25901 (47%)	23383 (49%)	0.04	8317 (51%)	0.07
Unknown[b]	4 (<0.1%)	0.00	6 (<0.1%)	8 (<0.1%)	0.01	2 (<0.1%)	0.00
Race							
White	18792 (38%)	0.00	20573 (38%)	15658 (33%)	0.10	3432 (21%)	0.36
Black/African American	10990 (22%)	0.04	11114 (20%)	9935 (21%)	0.01	4896 (30%)	0.23
Asian	2143 (4.3%)	0.05	2892 (5.3%)	2194 (4.6%)	0.03	478 (2.9%)	0.12
Other[b]	10057 (20%)	0.00	10893 (20%)	11077 (23%)	0.08	4577 (28%)	0.19
Unknown[b]	8034 (16%)	0.02	9257 (17%)	8776 (18%)	0.04	3035 (18%)	0.04
Hispanic ethnicity							
Yes	8944 (18%)	0.03	9085 (17%)	9822 (21%)	0.10	4384 (27%)	0.25
No	30320 (61%)	0.01	32938 (60%)	26378 (55%)	0.10	8179 (50%)	0.21
Unknown	10752 (21%)	0.04	12706 (23%)	11440 (24%)	0.02	3855 (23%)	0.01
Social Deprivation Index score	69 (32)	0.05	68 (33)	71 (32)	0.12	79 (27)	0.38
Unknown	1826 (3.7%)	-	2091 (3.8%)	1873 (3.9%)	-	649 (4.0%)	-
SDI quintile							
1	9878 (20%)	0.04	11646 (22%)	8691 (19%)	0.08	1801 (11%)	0.28
2	10321 (21%)	0.01	11420 (22%)	8994 (20%)	0.05	2522 (16%)	0.14
3	8865 (18%)	0.00	9665 (18%)	8360 (18%)	0.00	2921 (19%)	0.00
4	10043 (21%)	0.01	10688 (20%)	10158 (22%)	0.05	4088 (26%)	0.14
5	9083 (19%)	0.03	9219 (18%)	9564 (21%)	0.08	4437 (28%)	0.26
Body mass index (kg/m^2)	30.9 (6.9)	0.01	31.0 (6.7)	31.1 (6.8)	0.02	31.1 (7.2)	0.03
Unknown	6043 (12%)	-	6909 (13%)	6878 14%)	-	2830 (17%)	-
Systolic blood pressure (mmHg)	131 (14)	0.07	132 (14)	133 (15)	0.09	134 (16)	0.16
Unknown	3263 (6.5%)	-	3916	3941 (8.3%)	-	1599 (9.7%)	-
Diastolic blood pressure (mmHg)	75 (8)	0.01	75 (8)	76 (8)	0.05	78 (9)	0.31
Unknown	3268 (6.5%)	-	3921 (7.2%)	3948 (8.3%)	-	1600 (9.7%)	-
Estimated glomerular filtration rate (ml/min/1.73 m^2)[c]	72 (23)	0.01	73 (23)	73 (26)	0.02	80 (27)	0.28
Unknown	1108 (2.2%)	-	1383 (2.5%)	1307 (2.7%)	-	436 (2.7%)	-
Serum creatinine (mg/dL)	1.12 (0.95)	0.00	1.12 (0.91)	1.16 (0.96)	0.04	1.09 (0.84)	0.03
Unknown	1846 (3.7%)	-	2142 (3.9%)	2062 (4.3%)	-	750 (4.6%)	-
Comorbidities							
Hypertension uncomplicated	27231 (54%)	0.04	30774 (56%)	25241 (53%)	0.07	7656 (47%)	0.19

TABLE 1 (Continued)

Characteristic	39–46 mmol/mol (5.7–6.5%) N = 50,016	smd	48–57 mmol/mol (6.5–7.4%) N = 54,729	58–85 mmol/mol (7.5–9.9%) N = 47,640	smd	≥86 mmol/mol (≥10%) N = 16,418	smd
Obesity	12685 (25%)	0.04	12921 (24%)	11429 (24%)	0.01	3821 (23%)	0.01
Renal failure	7840 (16%)	0.01	8790 (16%)	8970 (19%)	0.07	2660 (16%)	0.00
Chronic pulmonary disease	7955 (16%)	0.07	7385 (13%)	6238 (13%)	0.01	2239 (14%)	0.00
Hypertension complicated	6696 (13%)	0.02	6999 (13%)	7245 (15%)	0.07	2420 (15%)	0.06
Cardiac arrhythmia	7671 (15%)	0.05	7418 (14%)	5892 (12%)	0.04	1803 (11%)	0.08
Hypothyroidism	7161 (14%)	0.06	6778 (12%)	5297 (11%)	0.04	1260 (7.7%)	0.15
Fluid and electrolyte disorders	4794 (9.6%)	0.01	5004 (9.1%)	5338 (11%)	0.07	2515 (15%)	0.20
Congestive heart failure	4874 (9.7%)	0.01	5112 (9.3%)	4982 (10%)	0.04	1588 (9.7%)	0.01
Depression	5236 (10%)	0.05	4845 (8.9%)	4626 (9.7%)	0.03	1829 (11%)	0.08
Outpatient medications							
Insulin	6203 (12%)	0.19	11289 (21%)	17997 (38%)	0.39	8408 (51%)	0.70
Metformin	17076 (34%)	0.26	25602 (47%)	22935 (48%)	0.03	7554 (46%)	0.02
Sulfonylureas	3231 (6.5%)	0.20	7290 (13%)	9654 (20%)	0.20	3008 (18%)	0.14
DPP4 inhibitors	4090 (8.2%)	0.20	8508 (16%)	9694 (20%)	0.13	3198 (19%)	0.11
SGLT2 inhibitors	1520 (3.0%)	0.19	4399 (8.0%)	5667 (12%)	0.14	1450 (8.8%)	0.03
GLP1 agonists	2358 (4.7%)	0.11	4284 (7.8%)	6105 (13%)	0.17	2125 (13%)	0.18
Thiazolidinediones	772 (1.5%)	0.07	1467 (2.7%)	1696 (3.6%)	0.05	565 (3.4%)	0.05
Antihypertensives	31667 (63%)	0.03	35326 (65%)	30384 (64%)	0.02	9397 (57%)	0.15
Aspirin	12483 (25%)	0.02	14031 (26%)	13572 (28%)	0.06	4542 (28%)	0.05
Statins	24849 (50%)	0.09	29753 (54%)	25802 (54%)	0.00	8051 (49%)	0.11
Immunosuppressants	2232 (4.5%)	0.03	2157 (3.9%)	2105 (4.4%)	0.02	467 (2.8%)	0.06
Antidepressants	7545 (15%)	0.06	7072 (13%)	6292 (13%)	0.01	2017 (12%)	0.02
Antipsychotics	2328 (4.7%)	0.05	1990 (3.6%)	1926 (4.0%)	0.02	796 (4.8%)	0.06
Health services utilization							
Baseline inpatient encounters	0.28 (0.80)	0.04	0.25 (0.76)	0.32 (0.87)	0.08	0.45 (1.03)	0.22
Baseline outpatient encounters[d]							
0	2053 (4.1%)	0.00	2308 (4.2%)	2639 (5.5%)	0.05	1994 (12%)	0.32
1–3	10139 (20%)	0.01	11319 (21%)	10262 (22%)	0.02	4504 (27%)	0.16
4–9	16576 (33%)	0.00	18250 (33%)	15072 (32%)	0.04	4674 (29%)	0.10
10+	21248 (43%)	0.01	22852 (42%)	19667 (41%)	0.01	5246 (32%)	0.20
Baseline ED encounters	0.42 (1.27)	0.05	0.35 (1.09)	0.45 (1.23)	0.08	0.70 (1.73)	0.26
Baseline outpatient medications	15 (26)	0.01	15 (27)	16 (29)	0.03	14 (26)	0.04

Note: Statistics presented: *n* (column %); mean (SD); Abbreviations: smd, standardized mean difference.

[a]Individual A1C measurements below 3 or above 20 excluded from A1C statistics.

[b]'Unknown' sex collapsed from categories 'Unknown' and 'Other'; 'Unknown' race collapsed from categories 'Unknown', 'Refuse' and 'No Information'; 'Other' race collapsed from categories 'Native Hawaiian/Pacific Islander', 'American Indian or Alaska Native', 'Other' and 'Multiple race'.

[c]Chronic Kidney Disease (CKD) Epidemiology Collaboration (CKD-EPI) equation without race coefficient for Black or African American patients (ref Diao et al.).

[d]Outpatient encounters comprise ambulatory visits (AVs) and other ambulatory encounters (OAs).

TABLE 2 Risk of COVID-19 hospitalization and in-hospital death by baseline HbA1c

Outcome	39–46 mmol/mol (5.7–6.4%) *N* = 50016	48–57 mmol/mol (6.5–7.4%) *N* = 54729	58–85 mmol/mol (7.5–9.9%) *N* = 47640	≥86 mmol/mol (≥10%) *N* = 16418
Primary analysis				
First COVID−19 hospitalization, *n*	490	507	612	267
Person-days of follow-up	4,564,072	4,995,935	4,335,516	1,489,889
Unadjusted rate /1000 person-days	3.22	3.04	4.23	5.38
Unadjusted HR (95% CI)	1.06 (0.93, 1.20)	Reference	1.39 (1.24, 1.56)	1.76 (1.52, 2.04)
Adjusted HR[a] (95% CI)	0.92 (0.82, 1.04)	Reference	1.19 (1.06, 1.34)	1.40 (1.19, 1.64)
Adjusted HR (using SuperLearner)[a,b]	0.98	Reference	1.19	1.36
In-hospital death from COVID−19, *n*	149	155	197	59
Person-days of follow-up	4,588,067	5,021,237	4,365,665	1,504,568
Unadjusted rate /person-days	0.97	0.92	1.35	1.18
Unadjusted HR (95%CI)	1.05 (0.84, 1.32)	Reference	1.46 (1.18, 1.80)	1.27 (0.94, 1.71)
Adjusted HR[a] (95% CI)	0.93 (0.75, 1.16)	Reference	1.29 (1.05, 1.59)	0.98 (0.70. 1.37)
Adjusted HR (using SuperLearner)[a,b]	1.03	Reference	1.27	1.01
Sensitivity analysis: Excluding patients with low predicted EHR continuity	*N = 5359*	*N = 5313*	*N = 5956*	*N = 2755*
First COVID−19 hospitalization, *n*	182	171	227	126
Person-days of follow-up	479,176	475,467	530,485	243,950
Unadjusted rate /1000 person-days	11.39	10.79	12.84	15.50
Unadjusted HR (95% CI)	1.06 (0.86, 1.30)	Reference	1.19 (0.97, 1.45)	1.43 (1.14, 1.80)
Adjusted HR[a] (95% CI)	0.91 (0.74, 1.11)	Reference	1.05 (0.86, 1.28)	1.49 (1.17, 1.91)
Adjusted HR (using SuperLearner)[a,b]	1.05	Reference	1.10	1.51
Exploratory analysis: Influenza outcome	*N* = 50016	*N* = 54729	*N* = 47640	*N* = 16418
First influenza hospitalization, *n*	104	105	118	51
Person-days of follow-up	4,593,603	5,027,377	4,374,286	1,506,669
Unadjusted rate /1000 person-days	0.68	0.63	0.81	1.02
Unadjusted HR (95%CI)	1.08 (0.83, 1.42)	Reference	1.29 (0.99, 1.68)	1.62 (1.16, 2.26)
Adjusted HR[a] (95% CI)	0.96 (0.74, 1.25)	Reference	1.08 (0.83, 1.39)	1.19 (0.83, 1.72)
Adjusted HR (using SuperLearner)[a,b]	1.00	Reference	1.14	1.17

[a]Adjusted for the following covariates: age, sex, race, Hispanic ethnicity, SDI quintile, BMI, systolic blood pressure, diastolic blood pressure, creatinine, baseline Elixhauser comorbidities (uncomplicated hypertension, obesity, renal failure, chronic pulmonary disease, complicated hypertension, cardiac arrhythmia, hypothyroidism, fluid and electrolyte disorders, congestive heart failure, depression), baseline medications (insulin, metformin, sulfonylurea, DPP4, SGLT2, GLP1, thiazolidinedione, antihypertensives, aspirin, statins, immunosuppressants, antidepressants, antipsychotics), baseline inpatient encounters, baseline outpatient encounters, baseline ED encounters and baseline outpatient medications.

[b]Propensity scores estimated using SuperLearner algorithm; 95% confidence intervals are not estimated.

with well-controlled HbA1c. Compared with patients with HbA1c 48–57 mmol/mol (6.5%–7.4%), the risk of COVID-19 hospitalizations was approximately 20% greater among those with HbA1c 58–85 (7.5%–9.9%) and 40% greater among those with HbA1c ≥86 (10%). The risk of in-hospital death from COVID-19 was higher among patients with HbA1c 58–85 (7.5%–9.9%) compared with those with HbA1c 48–57 (6.5%–7.4%), but the risk did not appear to be elevated in those with HbA1c ≥86 (10%). This could be due to limitations of our data, which do not capture deaths that occur outside the hospital setting. COVID-19 hospitalizations were accompanied by AKI in 40% of patients, and hypoglycaemia was also present in 5%–10% of patients. A co-diagnosis of ketoacidosis was also frequently observed in COVID-19 hospitalizations among patients with high HbA1c (7%–13% among those with HbA1c ≥58 mmol/mol [7.5%]), supporting previously reported case series that suggested ketoacidosis as a relatively common complication of COVID-19 infection.[34]

TABLE 3 Presence of diagnosis codes or laboratory values consistent with acute kidney injury, diabetic ketoacidosis and hypoglycaemia in COVID-19 hospitalizations and non-COVID-related hospitalizations by HbA1c group

Secondary outcome	Total	39–46 mmol/mol (5.7%–6.4%)	48–57 mmol/mol (6.5%–7.4%)	58–85 mmol/mol (7.5%–9.9%)	≥86 mmol/mol (≥10%)
COVID−19 hospitalizations, *n* (%)	2225	561	588	740	336
Acute kidney injury (diagnosis)	883 (39.7)	212 (37.8)	235 (40.0)	306 (41.4)	130 (38.7)
Acute kidney injury (lab)	865 (38.9)	216 (38.5)	235 (40.0)	283 (38.2)	131 (39.0)
Diabetic ketoacidosis	121 (5.4)	5 (0.9)	21 (3.6)	50 (6.8)	45 (13.4)
Hypoglycaemia (diagnosis)	206 (9.3)	31 (5.5)	56 (9.5)	78 (10.5)	41 (12.2)
Hypoglycaemia (lab)	133 (6.0)	30 (5.3)	36 (6.1)	50 (6.8)	17 (5.1)

4.2 | Discussion of previous literature

Previous studies have examined the association between diabetes and severity of COVID-19. Some studies specifically focused on patients with diabetes and whether severity of diabetes and its control was associated with COVID-19 outcomes. However, most of these studies included patients who were already admitted to the hospital for COVID-19 and were susceptible to selection bias (i.e. collider bias) where the results may not accurately reflect the association between diabetes severity or glucose control and outcomes of COVID-19 in the general diabetes population.

A population-based cohort study in the United Kingdom examined multiple risk factors for COVID-19 mortality among patients with type 1 or type 2 diabetes in a multivariable Cox proportional hazards model. Among many risk factors such as male gender, older age, renal impairment, non-white ethnicity, socioeconomic status, and cardiovascular disease, the authors found that type 2 diabetes patients with higher HbA1c (≥86 mmol/mol [10%]) had an increased risk of COVID-19 death compared with patients with HbA1c 48–53 (6.5%–7.0%) (HR 1.61, 95% CI 1.47–1.77), but this was evaluated without the use of propensity score weighting or matching. In our study, we did not find a significant association between very high HbA1c (≥86 mmol/mol [10%]) and in-hospital COVID-19 death. We could not capture deaths that occur outside the hospital, which may account for the difference in results. For example, if patients with poorly controlled HbA1c were less likely to present for care and more likely to die at home, an association between elevated HbA1c and increased mortality could be obscured. In another recent study of patients with diabetes who were tested for COVID-19 in Israel, HbA1c ≥75 mmol/mol (9%) was associated with an increased risk of COVID-19 hospitalization accounting for other baseline conditions (adjusted odds ratio 4.95, 95% CI 1.55, 15.76).[35] Because all study participants were tested for COVID-19, the rates of hospitalization cannot be directly compared with our study.

4.3 | COVID-19 hospitalization comorbidities

Among COVID-19 hospitalizations in our study, AKI was often present as a comorbid condition. Hypoglycaemia was more common among patients with higher HbA1c, likely due to more frequent use of insulin and other hypoglycaemic agents. The most notable characteristic of COVID-19 hospitalizations was that ketoacidosis diagnoses were very common in higher HbA1c groups. There have been previous reports that ketoacidosis may be a common complication of COVID-19.[19] Further investigation is needed on whether and how much this risk is attributable to type 1 diabetes patients, or if it is associated with certain antidiabetic drugs such as SGLT2 inhibitors.

4.4 | Comparison with influenza outcome

When we examined influenza as a 'negative control' outcome, we found no association between HbA1c level and influenza hospitalization. This could indicate that poor glucose control pre-infection is a more important risk factor for COVID-19 compared with influenza infection and that there may be a specific causal mechanism that explains the stronger association. However, due the low number of influenza hospitalizations during the study period, the power of this analysis is limited.

4.5 | Study limitations

Our study had several limitations. First, residual confounding may be present in our analyses. Although we used propensity score weighting and additional covariate adjustment to minimize this, perfect balance on the covariates among HbA1c comparison groups was not achieved in the primary analysis. When we used a machine learning algorithm to estimate propensity score weights, we were able to achieve better balance across covariates, and the effect estimates were similar to those in the primary analysis.

Although statistically reliable confidence intervals cannot be estimated using this method,[32] the sensitivity analysis allowed us to check that the effect estimates were robust to modelling choices. The capture of key covariates and confounders in the data were also limited and we had no ability to adjust for some potentially important variables such as diabetes duration and albumin to creatinine ratio. However, sensitivity analyses in which additional covariates were adjusted for – such as renin-angiotensin system antagonist use – showed no meaningful change in study results.[36] Second, it is possible that we missed COVID-19 deaths. Because of the limited scope of the CRN data, we were unable to capture in-home deaths or deaths at other healthcare facilities. During the height of the COVID-19 epidemic in New York City, many patients were unable to receive inpatient care and may have died in home or at a long-term care facility. Third, an important limitation of this research is that a large proportion of medical events occurring in New York City are not captured in any single EHR for each patient. A premise of this study was that the CRN would offer sufficiently complete data to support valid causal inference, an assumption that we are able to partly test by restriction of the data sample to patients with very high predicted data continuity in the CRN. That sensitivity analysis yielded essentially unchanged hazard ratios for the primary outcome and results that cohere with the limited existing evidence, although absolute event rates were markedly different. The implication is that, even in the highly fragmented New York City health system, the CRN is able to provide meaningful estimates of relative hazards, although assessment of absolute risks needs to be interpreted with great caution.

There are additional limitations with respect to generalisability. Our study used data from the early phase of the pandemic, and it is unclear whether similar findings will be present in the current phase. While the CRN data are comprised of a diverse patient population, it is limited to patients in New York City, and the patient selection process and analysis methods including propensity score weighting may have additionally selected a specific study sample in terms of patient characteristics and should be considered when interpreting our results. In addition, we excluded patients with a HbA1c <39 mmol/mol (5.7%) from the primary analysis, so this study is not informative regarding patients with extremely well-controlled diabetes.

5 | CONCLUSIONS

Higher HbA1c was associated with a greater risk of COVID-19 hospitalizations. While the causal determinants of increased risk of severe COVID-19 outcomes remain unclear, it appears prudent to prioritize HbA1c control and diabetes self-management education as possible means of reducing this risk and to focus efforts on promoting vaccination against COVID-19 in patients with poor HbA1c management.

ACKNOWLEDGEMENTS

This study was funded by the Patient Centered Outcomes Research Institute (PCORI) (CER-2017C3-9230) and the National Cancer Institute (P30-CA008748).

CONFLICTS OF INTEREST

The authors have no conflicts of interest to report.

REFERENCES

1. Atkins JL, Masoli JAH, Delgado J, et al. Preexisting comorbidities predicting COVID-19 and mortality in the UK Biobank Community Cohort. *J Gerontol A Biol Sci Med Sci*. 2020;75(11):2224-2230. doi:10.1093/gerona/glaa183
2. Barron E, Bakhai C, Kar P, et al. Associations of type 1 and type 2 diabetes with COVID-19-related mortality in England: a whole-population study. *Lancet Diabetes Endocrinol*. 2020;8(10):813-822. doi:10.1016/S2213-8587(20)30272-2
3. Dennis JM, Mateen BA, Sonabend R, et al. Type 2 diabetes and COVID-19–related mortality in the critical care setting: a national cohort study in England, March–July 2020. *Diabetes Care*. 2021;44(1):50-57. doi:10.2337/dc20-1444
4. Gregory JM, Slaughter JC, Duffus SH, et al. COVID-19 severity Is tripled in the diabetes community: a prospective analysis of the pandemic's impact in type 1 and type 2 diabetes. *Diabetes Care*. 2021;44(2):526-532. doi:10.2337/dc20-2260
5. McGurnaghan SJ, Weir A, Bishop J, et al. Risks of and risk factors for COVID-19 disease in people with diabetes: a cohort study of the total population of Scotland. *Lancet Diabetes Endocrinol*. 2021;9(2):82-93. doi:10.1016/S2213-8587(20)30405-8
6. You JH, Lee SA, Chun S-Y, et al. Clinical outcomes of COVID-19 Patients with type 2 diabetes: a population-based study in Korea. *Endocrinol Metab Seoul Korea*. 2020;35(4):901-908. doi:10.3803/EnM.2020.787
7. Carey IM, Critchley JA, DeWilde S, Harris T, Hosking FJ, Cook DG. Risk of infection in type 1 and type 2 diabetes compared with the general population: a matched cohort study. *Diabetes Care*. 2018;41(3):513-521. doi:10.2337/dc17-2131
8. Li W, Moore MJ, Vasilieva N, et al. Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*. 2003;426(6965):450-454. doi:10.1038/nature02145
9. Hoffmann M, Kleine-Weber H, Schroeder S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell*. 2020;181(2):271-280.e8. doi:10.1016/j.cell.2020.02.052
10. Holman N, Knighton P, Kar P, et al. Risk factors for COVID-19-related mortality in people with type 1 and type 2 diabetes in England: a population-based cohort study. *Lancet Diabetes Endocrinol*. 2020;8(10):823-833. doi:10.1016/S2213-8587(20)30271-0
11. Hayek AAA, Robert AA, Matar AB, et al. Risk factors for hospital admission among COVID-19 patients with diabetes. A

study from Saudi Arabia. *Saudi Med J*. 2020;41(10):1090-1097. doi:10.15537/smj.2020.10.25419

12. Cheng Y, Yue L, Wang Z, Zhang J, Xiang G. Hyperglycemia associated with lymphopenia and disease severity of COVID-19 in type 2 diabetes mellitus. *J Diabetes Complications*. 2021;35(2):107809. doi:10.1016/j.jdiacomp.2020.107809
13. Xu PP, Tian RH, Luo S, et al. Risk factors for adverse clinical outcomes with COVID-19 in China: a multicenter, retrospective, observational study. *Theranostics*. 2020;10(14):6372-6383. doi:10.7150/thno.46833
14. Zhang N, Wang C, Zhu F, et al. Risk factors for poor outcomes of diabetes patients with COVID-19: a single-center, retrospective study in early outbreak in China. *Front Endocrinol*. 2020;11:571037. doi:10.3389/fendo.2020.571037
15. Zhu L, She Z-G, Cheng XU, et al. Association of blood glucose control and outcomes in patients with COVID-19 and preexisting type 2 diabetes. *Cell Metab*. 2020;31(6):1068-1077.e3. doi:10.1016/j.cmet.2020.04.021
16. Kaushal R, Hripcsak G, Ascheim DD, et al. Changing the research landscape: the New York City Clinical Data Research Network. *J Am Med Inform Assoc JAMIA*. 2014;21(4):587-590. doi:10.1136/amiajnl-2014-002764
17. Our Data. INSIGHT Clinical Research Network. Accessed May 26, 2021. https://insightcrn.org/our-data/
18. Track Coronavirus Cases in Places Important to You: New York City metro area. The New York Times. Accessed May 26, 2021. https://www.nytimes.com/interactive/2020/us/covid-cases-deaths-tracker.html
19. Palermo NE, Sadhu AR, McDonnell ME. Diabetic ketoacidosis in COVID-19: unique concerns and considerations. *J Clin Endocrinol Metab*. 2020;105(8): doi:10.1210/clinem/dgaa360
20. Pal R, Banerjee M, Yadav U, Bhattacharjee S. Clinical profile and outcomes in COVID-19 patients with diabetic ketoacidosis: a systematic review of literature. *Diabetes Metab Syndr*. 2020;14(6):1563-1569. doi:10.1016/j.dsx.2020.08.015
21. Lopes JA, Jorge S. The RIFLE and AKIN classifications for acute kidney injury: a critical and comprehensive review. *Clin Kidney J*. 2013;6(1):8-14. doi:10.1093/ckj/sfs160
22. Bonds DE, Miller ME, Bergenstal RM, et al. The association between symptomatic, severe hypoglycaemia and mortality in type 2 diabetes: retrospective epidemiological analysis of the ACCORD study. *BMJ*. 2010;340:b4909.
23. International Hypoglycaemia Study Group. Glucose concentrations of less than 3.0 mmol/L (54 mg/dL) should be reported in clinical trials: a joint position statement of the American Diabetes Association and the European Association for the study of diabetes. *Diabetes Care*. 2017;40(1):155-157. doi:10.2337/dc16-2215
24. International Hypoglycaemia Study Group. Minimizing hypoglycemia in diabetes. *Diabetes Care*. 2015;38(8):1583-1591. doi:10.2337/dc15-0279
25. Zhang Y, Ancker JS, Hall J, Khullar D, Wu Y, Kaushal R. Association between residential neighborhood social conditions and health care utilization and costs. *Med Care*. 2020;58(7):586-593. doi:10.1097/MLR.0000000000001337
26. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8-27. doi:10.1097/00005650-199801000-00004
27. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43(11):1130-1139. doi:10.1097/01.mlr.0000182534.19832.83
28. van Walraven C, Austin PC, Jennings A, Quan H, Forster AJ. A modification of the Elixhauser comorbidity measures into a point system for hospital death using administrative data. *Med Care*. 2009;47(6):626-633. doi:10.1097/MLR.0b013e31819432e5
29. Patterson JA, Stuart EA, Ford JB. Use of propensity score methods to address adverse events associated with the storage time of blood in an obstetric population: a comparison of methods. *BMC Res Notes*. 2016;9:367. doi:10.1186/s13104-016-2169-1
30. van Buuren S. *Flexible Imputation of Missing Data*, 2nd ed. CRC Press, Taylor & Francis Group; 2018.
31. van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Stat Appl Genet Mol Biol*. 2007;6:Article25. doi:10.2202/1544-6115.1309
32. Zivich PN, Breskin A. Machine learning for causal inference: on the use of cross-fit estimators. *Epidemiology*. 2021;32(3):393-401.
33. Lin KJ, Singer DE, Glynn RJ, Murphy SN, Lii J, Schneeweiss S. Identifying patients with high data completeness to improve validity of comparative effectiveness research in electronic health records data. *Clin Pharmacol Ther*. 2018;103(5):899-905. doi:10.1002/cpt.861
34. Papadopoulos VP, Koutroulos MV, Zikoudi DG, et al. Diabetes-related acute metabolic emergencies in COVID-19 patients: a systematic review and meta-analysis. *Diabetol Int*. 2021;12(4):445-459. doi:10.1007/s13340-021-00502-9
35. Merzon E, Green I, Shpigelman M, et al. Haemoglobin A1c is a predictor of COVID-19 severity in patients with diabetes. *Diabetes Metab Res Rev*. 2021;37(5):e3398. doi:10.1002/dmrr.3398
36. Davis TME, Davis WA. Influence of renin-angiotensin system inhibitors on lower-respiratory tract infections in type 2 diabetes: the Fremantle diabetes study phase II. *Diabetes Care*. 2020;43(9):2113-2120. doi:10.2337/dc20-0895

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of the article at the publisher's website.

How to cite this article: Min JY, Williams N, Simmons W, et al. Baseline haemoglobin A1c and the risk of COVID-19 hospitalization among patients with diabetes in the INSIGHT Clinical Research Network. *Diabet Med*. 2022;39:e14815. doi:10.1111/dme.14815