

# Adverse outcomes in COVID-19 and diabetes: a retrospective cohort study from three London teaching hospitals

Chioma Izzi-Engbeaya <sup>1,2</sup>, Walter Distaso,<sup>3</sup> Anjali Amin,<sup>1</sup> Wei Yang,<sup>4</sup> Oluwagbemiga Idowu,<sup>1</sup> Julia S Kenkre,<sup>4</sup> Ronak J Shah,<sup>5</sup> Evelina Woin,<sup>5</sup> Christine Shi <sup>1</sup>, Nael Alavi,<sup>5</sup> Hala Bedri,<sup>5</sup> Niamh Brady,<sup>5</sup> Sophie Blackburn,<sup>5</sup> Martina Leczycka,<sup>5</sup> Sanya Patel,<sup>1</sup> Elizaveta Sokol,<sup>5</sup> Edward Toke-Bjolgerud,<sup>5</sup> Ambreen Qayum,<sup>1</sup> Mariana Abdel-Malek,<sup>4</sup> David C D Hope,<sup>2</sup> Nick S Oliver,<sup>1,2</sup> Vasiliki Bravis,<sup>1</sup> Shivani Misra,<sup>1,2</sup> Tricia M Tan <sup>2,4</sup>, Neil E Hill <sup>1,2</sup>, Victoria Salem <sup>1,2</sup>

**To cite:** Izzi-Engbeaya C, Distaso W, Amin A, *et al*. Adverse outcomes in COVID-19 and diabetes: a retrospective cohort study from three London teaching hospitals. *BMJ Open Diab Res Care* 2021;**9**:e001858. doi:10.1136/bmjdr-2020-001858

► Supplemental material for this paper are available online. To view these files, please visit the journal online (<http://dx.doi.org/10.1136/bmjdr-2020-001858>).

CI-E and WD are joint first authors.  
NEH and VS are joint senior authors.

Received 25 August 2020  
Revised 23 November 2020  
Accepted 25 November 2020



© Author(s) (or their employer(s)) 2021. Re-use permitted under CC BY. Published by BMJ.

For numbered affiliations see end of article.

**Correspondence to**  
Dr Victoria Salem;  
[v.salem@imperial.ac.uk](mailto:v.salem@imperial.ac.uk)

## ABSTRACT

**Introduction** Patients with diabetes mellitus admitted to hospital with COVID-19 have poorer outcomes. However, the drivers of poorer outcomes are not fully elucidated. We performed detailed characterization of patients with COVID-19 to determine the clinical and biochemical factors that may be drivers of poorer outcomes.

**Research design and methods** This is a retrospective cohort study of 889 consecutive inpatients diagnosed with COVID-19 between March 9 and April 22, 2020 in a large London National Health Service Trust. Unbiased multivariate logistic regression analysis was performed to determine variables that were independently and significantly associated with increased risk of death and/or intensive care unit (ICU) admission within 30 days of COVID-19 diagnosis.

**Results** 62% of patients in our cohort were of non-white ethnic background and the prevalence of diabetes was 38%. 323 (36%) patients met the primary outcome of death/admission to the ICU within 30 days of COVID-19 diagnosis. Male gender, lower platelet count, advancing age and higher Clinical Frailty Scale (CFS) score (but not diabetes) independently predicted poor outcomes on multivariate analysis. Antiplatelet medication was associated with a lower risk of death/ICU admission. Factors that were significantly and independently associated with poorer outcomes in patients with diabetes were coexisting ischemic heart disease, increasing age and lower platelet count.

**Conclusions** In this large study of a diverse patient population, comorbidity (ie, diabetes with ischemic heart disease; increasing CFS score in older patients) was a major determinant of poor outcomes with COVID-19. Antiplatelet medication should be evaluated in randomized clinical trials among high-risk patient groups.

## INTRODUCTION

In February 2020, the first cohort studies from Wuhan reported clinical outcomes of patients treated for COVID-19.<sup>1–5</sup> Patients with hypertension and ischemic heart disease (IHD) were over-represented in hospital or intensive

## Significance of this study

### What is already known about this subject?

- COVID-19 is a condition caused by infection with the novel coronavirus SARS-CoV-2.
- People with diabetes are over-represented in studies of patients with COVID-19, and people with diabetes appear to have poorer outcomes in these cohorts.

### What are the new findings?

- We demonstrate that in our large cohort of patients admitted with COVID-19 to three London teaching hospitals, diabetes is associated with increased risk of admission to intensive care unit (ICU) and/or death within 30 days of diagnosis of COVID-19 on univariate but not multivariate analysis.
- Within the entire cohort, unbiased multivariate logistic regression demonstrated that use of antiplatelet medication was associated with a reduced risk of ICU admission/death, while lower platelet count and higher Clinical Frailty Scale score were associated with an increased risk of ICU admission/death.
- Among patients with diabetes, unbiased multivariate logistic regression demonstrated that pre-existing ischemic heart disease was strongly associated with an increased risk of ICU admission/death and the protective effect of antiplatelet medication was lost.

### How might these results change the focus of research or clinical practice?

- These results will add to the body of knowledge used to inform shielding guidance for people with diabetes, especially those with coexisting ischemic heart disease.
- Prospective randomized controlled studies of the use of antiplatelet medication in COVID-19 are required to determine if antiplatelet medication improves outcomes in patients with COVID-19 and if this remains true for patients with established vascular complications of the metabolic syndrome.

care unit (ICU) admissions,<sup>1 3 4</sup> and some studies have reported that patients with these comorbidities are more likely to have poorer outcomes even after adjustment for age and smoking status.<sup>6</sup>

The prevalence of diabetes in patients admitted to ICU for severe COVID-19 (typically 20%–30%)<sup>3 5</sup> greatly exceeds the adult population prevalence of diabetes (ie, 6.8% in the UK).<sup>7</sup> Patients infected with the prior two human coronaviruses (severe acute respiratory syndrome coronavirus and Middle East respiratory syndrome coronavirus) were more likely to experience more severe disease and die if they had a diagnosis of diabetes.<sup>8 9</sup> However, granularity on the way diabetes interacts with the natural history of COVID-19 has been slower to come and remains an area of intense concern for patients with diabetes.<sup>10</sup>

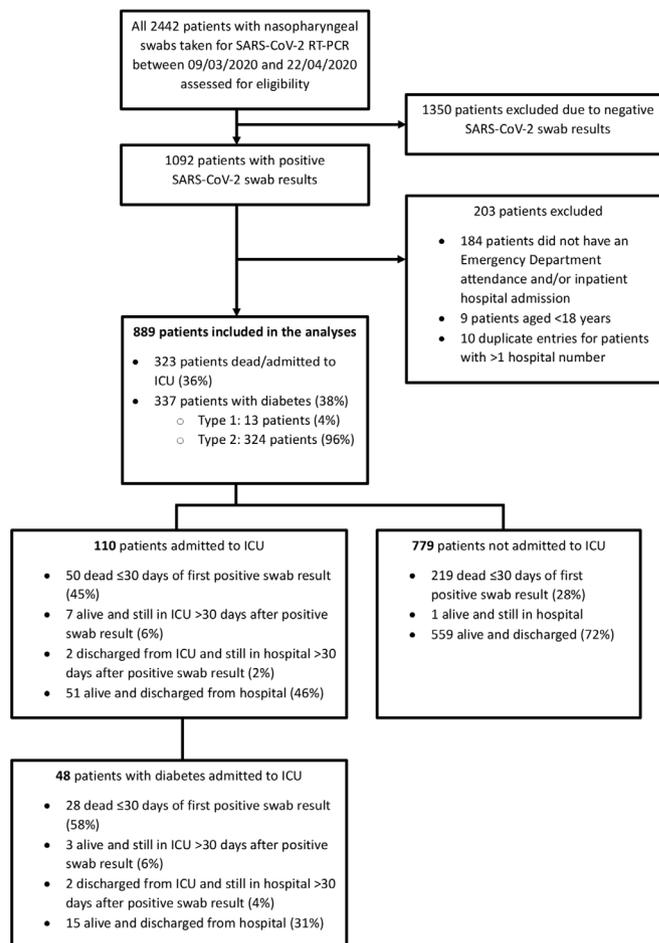
Two major hospital cohort studies provided early insight about the interaction between diabetes and COVID-19. Zhu *et al*<sup>11</sup> concluded that good glycemic control in the acute hospital setting in China was an important factor for better outcomes in patients with pre-existing type 2 diabetes, although it was difficult to ascertain whether this was confounded by the possibility that poorer glycemia in hospital was a marker of a more severe inflammatory response or the decision to use corticosteroids. The Coronavirus SARS-CoV-2 and Diabetes Outcomes (CORONADO) study reported that on multivariate analysis of patients with diabetes hospitalized for COVID-19, body mass index (BMI), but not glycated hemoglobin (HbA1c), was positively and independently associated with poorer outcomes.<sup>12</sup> More recently, a population-based study from England indicated that higher HbA1c increased the risk of mortality of patients with COVID-19 and diabetes, while the relationship between BMI and mortality was U-shaped.<sup>13</sup>

Here we provide a report of an unbiased/unselected multivariate analysis of all patients hospitalized with swab-positive COVID-19 in three London teaching hospitals, under the umbrella of Imperial College Healthcare NHS Trust (ICHNT), during the initial peak period of infections.

## RESEARCH DESIGN AND METHODS

### Study setting

We performed this retrospective cohort study at ICHNT, which includes three hospitals admitting patients with COVID-19 (Charing Cross Hospital, Hammersmith Hospital and St Mary's Hospital). All patients who had a nasopharyngeal swab taken to determine SARS-CoV-2 infection between March 9 and April 22, 2020 were identified on an automated search of the laboratories' computer systems. COVID-19 diagnosis was determined by the presence of SARS-CoV-2 infection as evidenced by a positive reverse transcription-PCR (RT-PCR) result from a nasopharyngeal swab. RT-PCR was performed on nasopharyngeal swabs by North West London Pathology



**Figure 1** All 2442 patients who had nasopharyngeal swabs taken for RT-PCR to detect the presence of SARS-CoV-2 (and diagnose COVID-19) between March 9 and April 22, 2020 were assessed for eligibility. 1350 patients with negative swab results were excluded. Of 1092 patients with positive swab results, 203 were excluded (as they did not have an emergency department attendance and/or inpatient hospital admission or they were aged less than 18 years), and 889 patients were included in the statistical analyses. ICU, intensive care unit; RT-PCR, reverse transcription-PCR.

staff in the laboratories in the constituent hospitals of ICHNT.

Excluded from data collection were all patients with negative SARS-CoV-2 swab results, patients who did not have an emergency department attendance and/or an inpatient admission within ICHNT, and all patients aged less than 18 years (figure 1). For patients with more than one positive swab result or multiple hospital admissions during the data collection period, the date of the earliest swab result and the most significant hospital admission in terms of COVID-19 severity, respectively, were used for data collection.

### Data collection

Hospital numbers of patients with positive swab results were checked against electronic health records held on the ICHNT computer system (Cerner Corporation, Kansas City, USA). Data are uploaded onto Cerner

manually by health practitioners and administrators to record and store patient notes, prescribe and dispense medication, and store investigation results and clinic letters. Furthermore, Cerner is connected to the National Health Service (NHS) Spine, which synchronizes other data (such as date of death).

Demographic, clinical and biochemical data (on the date of admission/emergency department attendance or on the date of the first positive SARS-CoV-2 swab result if the patient had been admitted for  $\geq 1$  week), as well as outcome data, were collected for all eligible patients up to and including June 20, 2020 to ensure there was a minimum of 30 days of follow-up after the date of COVID-19 diagnosis to eliminate the issue of censoring.

Data collected included demographic and clinical characteristics (age, gender, ethnicity, most recent body weight), specific comorbidities (including classification and duration of diabetes) and medications on admission. A score using the Clinical Frailty Scale (CFS)<sup>14</sup> was assigned to every patient. Hematological and biochemical data on the day of COVID-19 diagnosis included full blood count, biochemistry and arterial blood gas results. Average capillary blood glucose (CBG) values were recorded for the first 3 days following COVID-19 diagnosis and the most recent HbA1c result (within 6 months) was recorded.

### Statistical analysis

The composite primary outcome was admission to ICU or death within 30 days of COVID-19 diagnosis, and data on the primary outcome were collected for all patients. This was chosen to capture patients with severe and/or life-threatening COVID-19. Univariate and then unrestricted multivariate logistic regression analyses were performed to look for factors associated with an increased risk of the primary outcome. Unlike a Cox proportional hazards approach, this analysis focuses more on 'if' a patient gets severely ill rather than survival duration.

The first stage of analysis was descriptive. Quantitative data are expressed as mean $\pm$ SD. Categorical variables were given as percentage (number) of participants. Univariate logistic regression models were used to calculate OR associated with the primary outcome. Comparison between two groups was analyzed using Student's t-tests (normally distributed data) or Mann-Whitney U test (non-normally distributed data) for continuous variables. Comparison of categorical variables was analyzed by Fisher's exact test or  $\chi^2$  test. A difference with a two-sided  $\alpha < 0.05$  was considered statistically significant.

Multivariable logistic regression was applied to assess the independent association of the primary outcome with all the clinical and biological features included in the study. For the top level (unrestricted) analysis only data sets with  $< 5\%$  missing values were included, leaving a total number of 716 patients and 37 predictors (variables). No data points were imputed. Data were winsorized (at 1% level) to neutralize the effects of the most extreme values. Following this first step, a regularized regression

(smoothly clipped absolute deviation, SCAD) analysis<sup>15-17</sup> was performed to independently account for the overdimensionality of the unrestricted multivariate. This sensitivity analysis allows unbiased selection of the most relevant and important predictors and controlling the errors involved in variable selection, without the need to re-estimate the model in a stepwise fashion. Results are presented as the estimate parameter, the p value of the association and the marginal effect size.

Next, only patients with diabetes were included in multivariate and regularized regression analyses to look for elements of their clinical presentation that were associated with poorer outcome with COVID-19. Finally, marginal effects for the most significant (independent) risk factors were substratified by age (for the CFS) and the presence or absence of IHD or usage of antiplatelet agents (for platelet counts). All statistical analyses were performed on GraphPad Prism V.8.1 (descriptive statistics) and custom-written MaTLAB scripts (multivariate analyses). A detailed description of statistical methods can be found in the online supplemental material.

## RESULTS

### Study cohort

We report on 889 consecutive adult patients with a confirmed (SARS-CoV-2 swab-positive) diagnosis of COVID-19 admitted to three central London hospitals within the same NHS Trust (ICHNT) between March 9 and April 22, 2020. A total of 323 (36%) of these patients met the primary outcome of death or admission to the ICU within 30 days of diagnosis. In the cohort of patients studied here, 337 of 889 (38%) had a diagnosis of diabetes, 96% (324 of 337) of people with diabetes had type 2 diabetes (including 4 patients who were diagnosed during their admission for COVID-19) and the remainder (13 patients) had type 1 diabetes. The study flow diagram is shown in figure 1.

### Factors prior to hospital admission

The characteristics of the entire cohort are shown in online supplementary table S1. Of all patients treated for COVID-19 in our NHS Trust, the average ( $\pm$ SD) age was 65.8 ( $\pm 17.5$ ) years and 60% were men. The average ( $\pm$ SD) weight for men was 80.8 ( $\pm 19.8$ ) kg and for women 72.6 ( $\pm 24.2$ ) kg. Of our cohort, 62% were from non-white ethnic background, including 17% black and 11% South Asian. The most common comorbidities were hypertension (47%), hypercholesterolemia (43%), diabetes (38%) and IHD (16%).

By comparison, among the patients who had diabetes treated for COVID-19 (online supplementary table S2), the average age was 68.5 $\pm$ 14.6 years and 66% were men. The average weight was 83.5 ( $\pm 18.9$ ) kg for men and 77.8 ( $\pm 24.5$ ) kg for women. In this cohort of patients with diabetes, 72% were non-white, including 20% black and 14% South Asian. Other metabolic comorbidities were much more common in patients with diabetes compared

with the whole cohort: a medical history of hypertension was present in 70%, IHD in 28% and hyperlipidemia in 51%. The average HbA1c was 64.7 ( $\pm 21.7$ ) mmol/mol (8%), 27% were taking an ACE inhibitor and 63% were taking a statin.

### Factors at the time of diagnosis of COVID-19

Online supplemental table 3 summarizes the differences in presenting clinical characteristics between patients who fulfilled the primary outcome compared with patients who survived and were not admitted to ICU. Online supplemental table S4 compares the presentation features between patients with a diagnosis of diabetes and those without diabetes. Online supplemental table S5 shows the difference in presentation characteristics for patients with diabetes stratified by the primary outcome compared with patients with diabetes who survived and were not admitted to ICU. Biochemical parameters established to be associated with adverse outcomes, including procalcitonin and D-dimers, were significantly higher in the group who fulfilled the primary outcome of death/ICU admission (online supplemental table S3). The mean values for procalcitonin were slightly lower in patients with diabetes compared with those without diabetes (online supplemental table S4). However, among patients with diabetes, higher procalcitonin levels were still a predictor of poorer outcomes (online supplemental table S5). For both patients with diabetes and those without, a CBG below 10 mmol/L at presentation was associated with a lower risk of poor outcome. In patients without diabetes, this association was sustained after correction for age and gender; however, this was not sustained on multivariate analysis in patients with diabetes (see table 1).

Compared with patients who did not have diabetes, patients with diabetes had significantly lower arterial blood gas pH values, lower bicarbonate levels and higher serum potassium levels (although still, on average, within the laboratory reference range) at the time of COVID-19 diagnosis (online supplemental table S4). This hints at the possibility that patients with diabetes were more likely to have diabetic metabolic emergencies at presentation. Indeed, in our cohort there were 59 episodes of diabetic ketoacidosis associated with COVID-19 during the data collection period. However, admission measurements of acidemia were not significantly lower in patients with diabetes who met the primary outcome compared with patients who survived and were not admitted to ICU (online supplemental table S5).

Chest radiographs on diagnosis were reported by inhouse radiologists on a 6-point severity scale (0=normal, 5=widespread, dense bilateral infiltrates). Of the patients who survived without ICU admission, 60% had the classic patchy ground glass changes, with a median severity score of 2 out of 5. Among those who died/were admitted to ICU, 72% had the classic radiological features, with a median severity score of 3. These proportions were similar among the patients with diabetes. The frequency distribution of chest X-ray severity score by primary outcome,

with poorer outcomes associated with higher chest X-ray severity scores, was highly significantly different on  $\chi^2$  testing with 5 df ( $p < 0.0001$ ).

### Clinical features of the whole cohort associated with the primary outcome (death or ICU admission within 30 days of COVID-19 diagnosis)

On univariate analysis, the risk of the primary outcome was significantly higher in men and older age groups (online supplemental table S1). Hypertension, IHD, pre-existing renal impairment and heart failure were each associated with an increased risk of poor outcome on univariate analysis. Patients taking an ACE inhibitor had an OR of 1.57 (CI 1.10 to 2.22,  $p = 0.013$ ) for the primary outcome, but an increased risk was not statistically evident with patients on angiotensin receptor blocker drugs. Furthermore, the CFS score (ranging from 1 (very fit) to 9 (terminally ill)) was a good predictor of poorer outcomes, with patients scoring  $\geq 7$  having an OR of 2.34 (1.46 to 3.76,  $p = 0.0004$ ) for death/ICU admission compared with those scoring 1–2. Weight  $> 90$  kg did not reach significance for predicting poorer outcomes referenced against patients who weighed between 60 kg and 90 kg. However, patients weighing less than 60 kg were more likely to have poorer outcomes (online supplemental table S1).

After removing data sets where  $> 5\%$  data points were missing (ie, 170 patients), we ran an unselected multivariate logistic regression analysis of 37 regressors (variables) against the primary outcome on 719 patients (table 2). Age and gender retained significance as independent variables. As well as reporting  $p$  values for the significance of the independent contribution of a given variable, we have also calculated marginal effect sizes. This percentage value is a measure of the altered risk of primary outcome (death/ICU admission) as the given variable changes. The marginal effect of male gender was +9.3%; that is, all other factors being the same, the risk of primary outcome was 9.3% higher for men than for women. The CFS score was also found to be a significant independent predictor of death/ICU admission. Each unit rise in score (from 1 to 9) was associated with a 3.1% increased risk of death/ICU admission, all other things being equal.

The association between ethnicity and poor outcome was not sustained on multivariate analysis in our cohort. None of the individual comorbidities we collected data on (ie, diabetes, hypertension, hyperlipidemia, IHD, cerebrovascular disease, heart failure, chronic obstructive airway disease and cancer) were found to be independent predictors of mortality/ICU admission in our cohort, including diabetes.

Taking an antiplatelet drug was significantly and independently associated with a 10% lower risk of death and/or ICU admission (table 2). Related to this, platelet count was a significant predictor of poor outcome. Our data suggest that every reduction in platelet count of  $100 \times 10^9/L$  is associated with a 10% risk of death/ICU

**Table 1** Unbiased multivariate logistic regression analysis of 42 regressors (variables) against the primary outcome of death/ICU admission within 30 days of COVID-19 diagnosis in patients with diabetes mellitus (n=268 patients)

Regressor	Estimate	SE	P value	Marginal effect (%)
Age	0.059	0.019	<b>0.001</b>	<b>1.3</b>
Male gender	0.827	0.453	0.068	17.9
Ethnicity: South Asian	-1.352	0.736	0.066	-29.2
Ethnicity: black	0.216	0.578	0.709	4.7
Ethnicity: white	-0.091	0.508	0.858	-2.0
Type 1 diabetes	0.993	1.135	0.382	21.5
Active foot disease	0.075	0.839	0.929	1.6
Stroke	0.338	0.523	0.519	7.3
Hyperlipidemia	-0.503	0.447	0.261	-10.9
Ischemic heart disease	1.547	0.569	<b>0.007</b>	<b>33.4</b>
Heart failure	-0.956	0.613	0.119	-20.7
Hypertension	-0.101	0.504	0.841	-2.2
Chronic obstructive pulmonary disease	-0.499	0.661	0.450	-10.8
Active cancer	-0.383	0.727	0.598	-8.3
Insulin	0.228	0.992	0.818	4.9
GLP-1 receptor agonist	1.040	1.755	0.553	22.5
Metformin	-0.064	0.918	0.944	-1.4
Sulfonylurea	0.475	0.991	0.632	10.3
Dipeptidyl peptidase-IV inhibitor	-0.825	0.935	0.378	-17.8
Total number of medications for diabetes	0.137	0.800	0.864	3.0
Statin	-0.105	0.504	0.835	-2.3
ACE inhibitor	0.536	0.513	0.297	11.6
Angiotensin II receptor blocker	0.200	0.529	0.705	4.3
Antiplatelet drug	-0.735	0.486	0.130	-15.9
White cell count	-0.112	0.146	0.442	-2.4
Hemoglobin	-0.018	0.012	0.130	-0.4
Platelet count	-0.009	0.003	<b>0.001</b>	<b>-0.2</b>
Neutrophils	0.123	0.140	0.376	2.7
Lymphocytes	0.093	0.406	0.819	2.0
Serum sodium	0.077	0.039	<b>0.050</b>	<b>1.7</b>
eGFR on diagnosis	-0.012	0.010	0.224	-0.3
C reactive protein	0.002	0.002	0.536	0.0
Capillary blood glucose on diagnosis	0.034	0.038	0.369	0.7
Temperature	0.017	0.206	0.935	0.4
Respiratory rate on diagnosis	0.047	0.041	0.252	1.0
Heart rate on diagnosis	-0.004	0.014	0.774	-0.1
Systolic blood pressure	0.003	0.010	0.793	0.1
Diastolic blood pressure	0.015	0.017	0.391	0.3
NEWS on diagnosis	-0.001	0.110	0.990	0.0
Inspired oxygen delivered on diagnosis	0.000	0.011	0.974	0.0
Oxygen saturations on diagnosis	0.020	0.036	0.575	0.4
Maximum inspired oxygen required	0.072	0.012	<b>&lt;0.0001</b>	<b>1.6</b>

This is an unselected multivariate logistic (logit) analysis of all variables that were collected for patients admitted with swab-positive COVID-19 who had diabetes mellitus, as applied to the primary outcome of death or ICU admission within 30 days. 268 patients are included with 42 variables, with the only exclusions being those patients/variables for which  $\geq 5\%$  data points were unknown. For this reason, HbA1c is not included as a regressor as it would have reduced the number of patients included to 168, although of note in that regression HbA1c did not survive multiple correction (and neither did body weight). For categorical variables, a positive 'estimate' indicates an increased risk of the primary outcome (death or ICU admission) with that variable present, and a negative estimate indicates a reduced risk of the primary outcome if that variable is present. The p value is a measure of the confidence of that variable being an independent predictor of the primary outcome corrected for all of the other regressors listed. For continuous variables, a positive 'estimate' indicates an increasing risk of the primary outcome as the variable increases. Since in logistic regressions estimated coefficients cannot be interpreted as a measure of the contribution of the effect, we have also calculated marginal effects along with their SEs. A positive marginal effect indicates that an increase in that variable is associated with a fully adjusted increased risk of the primary outcome. The converse applies for negative marginal effects. For categorical variables, the marginal effect indicates the percentage increased risk of the primary outcome, if that variable exists. So, for example, patients with diabetes (all other things being equal) have a 33% increased risk of death/ICU if they have IHD. Statistically significant P values are shown in bold.

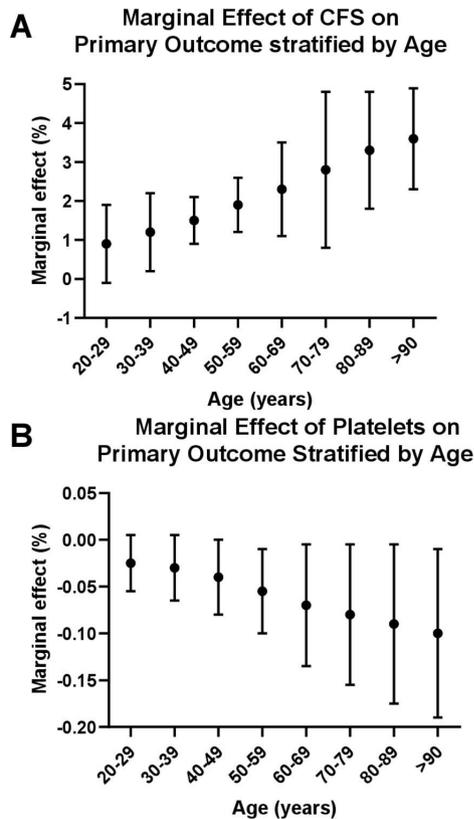
eGFR, estimated glomerular filtration rate; GLP-1, glucagon-like peptide-1; HbA1c, glycated hemoglobin; ICU, intensive care unit; IHD, ischemic heart disease; NEWS, National Early Warning Score.

**Table 2** Unbiased multivariate logistic regression analysis of 37 regressors (variables) against the primary outcome of death/ICU admission within 30 days of COVID-19 diagnosis (n=719 patients)

Regressor	Estimate	SE	P value	Marginal effect (%)
Age	0.029	0.011	<b>0.005</b>	<b>0.5</b>
Male gender	0.553	0.256	<b>0.031</b>	<b>9.3</b>
Ethnicity: South Asian	-0.321	0.441	0.466	-5.4
Ethnicity: black	-0.296	0.360	0.411	-5.0
Ethnicity: white	-0.179	0.272	0.510	-3.0
Diabetes mellitus	0.099	0.265	0.709	1.7
Stroke	-0.457	0.348	0.189	-7.7
Hyperlipidemia	0.117	0.277	0.673	2.0
Ischemic heart disease	0.596	0.364	0.102	10.1
Heart failure	0.046	0.408	0.910	0.8
Hypertension	0.207	0.285	0.468	3.5
Chronic obstructive pulmonary disease	-0.051	0.384	0.893	-0.9
Active cancer	-0.135	0.428	0.752	-2.3
Statin	-0.310	0.288	0.282	-5.2
ACE inhibitor	0.351	0.329	0.285	5.9
Angiotensin II receptor blocker	0.145	0.366	0.691	2.5
Antiplatelet drug	-0.616	0.319	<b>0.053</b>	<b>-10.4</b>
Clinical Frailty Scale score	0.183	0.078	<b>0.019</b>	<b>3.1</b>
White cell count	0.031	0.085	0.720	0.5
Hemoglobin	-0.008	0.006	0.182	-0.1
Platelet count	-0.004	0.001	<b>0.001</b>	<b>-0.1</b>
Neutrophils	0.009	0.086	0.920	0.1
Lymphocytes	-0.011	0.229	0.963	-0.2
Serum sodium	0.072	0.023	<b>0.002*</b>	<b>1.2</b>
Serum potassium	0.202	0.194	0.300	3.4
eGFR on diagnosis	-0.006	0.005	0.287	-0.1
C reactive protein	0.002	0.001	0.132	0.0
Temperature	-0.012	0.113	0.913	-0.2
Respiratory rate	0.034	0.021	0.103	0.6
Heart rate	0.006	0.008	0.453	0.1
Systolic blood pressure	0.010	0.006	0.088	0.2
Diastolic blood pressure	-0.002	0.010	0.831	0.0
National Early Warning Score	-0.015	0.055	0.790	-0.2
Inspired oxygen delivered on diagnosis	-0.012	0.006	<b>0.046</b>	<b>-0.2</b>
Oxygen saturations on diagnosis	-0.032	0.018	0.079	-0.5
Maximum inspired oxygen required during admission	0.068	0.007	<b>&lt;0.0001</b>	<b>1.2</b>

This is an unselected multivariate logistic (logit) analysis of all variables that were collected for patients admitted with swab-positive COVID-19. 719 patients are included with 37 variables, with the only exclusions being those patients/variables for which  $\geq 5\%$  data points were unknown. For categorical variables, a positive 'estimate' indicates an increased risk of the primary outcome (death or ICU admission) with that variable present, and a negative estimate indicates a reduced risk of the primary outcome if that variable is present. The p value is a measure of the confidence of that given variable being an independent predictor of the primary outcome corrected for all of the other regressors listed. For continuous variables, a positive 'estimate' indicates an increasing risk of the primary outcome as the variable increases. Since in logistic regressions estimated coefficients cannot be interpreted as a measure of the contribution of the effect, we have also calculated marginal effects. A positive marginal effect indicates that an increase in that variable is associated with a fully adjusted increased risk of the primary outcome. The converse applies for negative marginal effects. For categorical variables, the marginal effect indicates the percentage increased risk of the primary outcome, if that variable exists. So, for example, all other variables being equal, men have a 9.3% increased risk of death or ICU admission than women. Statistically significant P values are shown in bold.

\*The effect of serum sodium was skewed by patients with serum sodium  $>145$  mmol/L.  
eGFR, estimated glomerular filtration rate; ICU, intensive care unit.



**Figure 2** On unbiased univariate analysis of 798 patients admitted with COVID-19, the Clinical Frailty Scale (CFS) score was a significant independent predictor of poor outcomes, with a marginal effect size of 3% for every increase in score between 1 and 9 (see table 2). We stratified the marginal effect by age bands. There was a trend (A) for the marginal effect of the CFS score on the primary outcome to be greater in the older population. Similarly, there was a trend for the independent marginal effect of platelet count on the primary outcome to increase with age (B), but there was no interaction with the marginal effect size of platelet count on outcomes as stratified by the use of antiplatelet drugs (data not shown).

admission. This association remained after removal of patients with platelet counts below  $50 \times 10^9/L$  (the majority of whom had active, often hematological, malignancies). On univariate analysis, patients taking antiplatelet agents did not have different outcomes from those who were not taking antiplatelet agents (online supplemental table S1), although this is likely due to the fact that any beneficial effects of the drug are masked by their particular usage in high-risk groups, for example, those with IHD. Among patients with a pre-existing diagnosis of IHD, 51% of those who were *not* taking an antiplatelet drug died or were admitted to the ICU, compared with 53% of those with IHD who were taking an antiplatelet agent. Furthermore, in patients with diabetes, taking an antiplatelet agent was associated with a negative parameter estimate that was not statistically significant (table 1). Taken together, this suggests that antiplatelet agents may not protect as well against the platelet-consuming coagulopathy of severe disease in particularly high-risk groups.

Higher serum sodium at the time of COVID-19 diagnosis was also found to be a strong independent predictor of poor outcome; however, this association disappeared after removing patients with high sodium (ie,  $>145$  mmol/L). As a sensitivity analysis, we performed a regularized regression (SCAD), which selects the truly significant variables in the multivariate regression without the limitations of a stepwise approach (see online supplemental material). This SCAD analysis identified 23 variables driving the variance (online supplemental table S6). It supported the findings of the unrestricted multivariate regression (table 2).

### Clinical features of patients with diabetes associated with the primary outcome

Unrestricted multivariate logistic analysis with  $>95\%$  complete data sets was possible on 278 patients with diabetes (table 1). This excluded HbA1c from the analysis due to inadequate data. However, multivariate analysis on 197 patients for HbA1c did not produce a significant interaction with the primary outcome ( $p>0.9$ ) and neither was it a significant factor on univariate analysis (online supplemental table S2). Of all the associated comorbidities examined, IHD had a highly significant 33% marginal effect for increased likelihood of death/ICU admission for patients with diabetes. IHD was present as a comorbidity in 28% of patients with diabetes compared with 16% of the whole cohort. On univariate analysis, IHD was associated with a relative risk of 1.2 ( $p=0.02$ ) for poor outcome in patients with diabetes compared with patients with diabetes who did not have IHD (online supplemental table S2). Conversely, in all patients with IHD, outcomes were already poor (51% of patients with pre-existing IHD died or were admitted to ICU), and this was not further impacted on by the additional diagnosis of diabetes. Taken together this suggests that it is not diabetes per se, but its association with other cardiometabolic conditions (particularly established IHD) that confers risk of poorer outcomes with COVID-19. Other factors significantly and independently associated with poor outcome in patients with diabetes were age, lower platelet count and maximum inspired oxygen required during admission (table 1). In patients with diabetes, male gender was no longer a significant independent predictor of poorer outcome, although there was a trend for this to remain the case ( $p=0.07$ ). This was corroborated on SCAD analysis (online supplemental table S7).

### Stratified marginal effect size for the significant predictors that survived multiple logistic regression: platelet count and CFS score

We stratified the marginal effect of CFS score as an independent predictor of poor outcome in COVID-19 by age group (figure 2A). This confirms that CFS score is a useful predictor of the primary outcome in older age groups. The independent marginal effect of platelet count tends to increase with age (figure 2B), but there

was no interaction with the use of antiplatelet drugs ( $p>0.9$ ).

## DISCUSSION AND CONCLUSIONS

Here we report on a large cohort of patients diagnosed with COVID-19 over a 6-week period in a multisite NHS Trust in London. In our cohort, multivariate analysis revealed male gender, increased age, increased frailty and lower platelet count were independently associated with increased risk of ICU admission and/or death within 30 days of COVID-19 diagnosis, while taking antiplatelet medication was associated with a lower risk of poor outcome. Within the subset of our cohort with diabetes (96% type 2 and 4% type 1), multivariate analysis demonstrated pre-existing IHD, advancing age and lower platelet count were associated with increased risk of ICU admission and/or death.

Prior to the COVID-19 pandemic, data from the National Diabetes Inpatient Audit indicate the prevalence of diabetes among hospital inpatients in England and Wales was 18%,<sup>18</sup> reflecting a generalized increased risk of hospitalization among people with diabetes. Using data from a UK primary care database, Barron *et al*<sup>19</sup> reported that people with both type 1 and type 2 diabetes had multiply adjusted increased odds of dying in hospital with COVID-19 compared with those without diabetes. Our results show that people with diabetes are at increased risk of severe or life-threatening COVID-19, although this was driven by its tendency to coexist with other conditions, particularly IHD. The relative proportions of patients admitted with COVID-19 with type 1 and type 2 diabetes were similar to the population we serve, suggesting no difference in susceptibility based on diabetes type.

Hyperglycemia is a modifiable factor that may influence outcome in COVID-19, especially in people with diabetes. In our cohort, recent HbA1c was not a significant predictor of poor outcome, which is similar to some cohort studies<sup>12</sup> but not others.<sup>13</sup> In patients with diabetes (and in patients without diabetes), our data demonstrate that blood glucose  $\geq 10$  mmol/L (at the time of COVID-19 diagnosis and on average during the 72 hours following COVID-19 diagnosis) is associated with increased risk of death/ICU admission. Similar findings have been reported by groups from England,<sup>13</sup> France<sup>12</sup> and China.<sup>11 20</sup> This association is not sustained on multivariate analysis in our patients with diabetes and there is no prospective study to address whether maintaining blood glucose  $< 10$  mmol/L would improve outcomes for patients with diabetes with COVID-19. Furthermore, maintaining blood glucose  $< 10$  mmol/L may become even more challenging to achieve as the use of dexamethasone in the management of severe COVID-19 becomes more widespread, following publication of beneficial reports of its use in this context.<sup>21</sup>

It is now well established that patients with more severe manifestations of COVID-19 (including those that

need to be escalated to ICU care and those that die) are much more likely to have a diagnosis of diabetes than those who are documented as having a mild form of the infection.<sup>2 19 22</sup> Some people have recommended that patients with diabetes need to be more actively shielded, and diabetes may be associated with a higher risk of viral infection.<sup>23</sup> Evidence of increased risk of contracting COVID-19 in people with diabetes is lacking, with similar adult prevalence of diabetes (10.1%) and prevalence of diabetes in patients with COVID-19 (10.9%) reported by the Centers for Disease Control in America.<sup>24</sup> However, it is important to note that here we show that diabetes alone is not a major factor contributing to the risk of death/ICU admission, but rather its association with other consequences of the metabolic syndrome, particularly IHD, confers a higher risk of a poor outcome. For instance, patients (in our cohort) with diabetes have a 33% increased risk of death/ICU admission if they also have IHD.

In line with our observations that it is multimorbidity per se and not any single particular diagnosis that confers a strong increased risk of poorer outcomes in COVID-19 infection, we show that the CFS score is a robust, independent predictor of poor outcome on multivariate analysis. This is consistent with (COVID-19 and non-COVID-19) studies that show that as the CFS score increases, the likelihood of mortality increases.<sup>25 26</sup> However, we have also shown that the CFS score is much less useful in younger age groups. We chose death or ICU admission as our primary outcome measure as this incorporated all patients with severe/life-threatening COVID-19. Our cohort also included those who, with unfavorable chances of tolerating and surviving invasive ventilation, would not have been admitted to ICU due to pre-existing multimorbidity.

We report a strongly significant and independent risk of death/ICU admission as platelet count at presentation decreases. The inverse relationship between platelet count and risk of death/ICU admission with COVID-19 has also been reported by several other studies.<sup>22 27 28</sup> As microvascular and macrovascular thrombosis is increasingly being reported as a feature of severe COVID-19,<sup>29-31</sup> reduced platelet count may reflect consumptive coagulopathy (D-dimers were significantly higher in the primary outcome group), possibly in conjunction with direct effects of the virus on thrombopoiesis or platelet survival. We note that patients with coexisting diabetes and IHD are at particular risk of poorer outcomes with COVID-19, and that in these patients the protective benefits of antiplatelet usage are abrogated. Differing recommendations about the use of anticoagulants in hospitalized patients with COVID-19 have been made in recently published international guidelines,<sup>32</sup> while several prospective randomized trials evaluating the effects of anticoagulation on COVID-19 mortality are currently underway.

Strengths of our study include a diverse population of patients and indepth characterization, including CFS

and numerous prehospital and presentation factors. We selected a statistical approach focused on a primary outcome measure with no selection bias for the multivariate analysis that produced intuitive results, which survived robust sensitivity analysis.

Large population studies (of millions of patients), using primary care databases in England, have reported increased mortality risk among patients with COVID-19 who have diabetes.<sup>19 33</sup> Similarly, several meta-analyses (of thousands of patients) have reported increased risk of poor outcomes and/or death in people with COVID-19 who have diabetes.<sup>34 35</sup> In our cohort, diabetes per se was not associated with an increased risk of ICU admission and/or death on multivariate analysis (which included 37 variables including presenting clinical features as well as pre-existing comorbidities). It is clear that patients with diabetes are over-represented among those that have the poorest outcomes with COVID-19, and our findings suggest that this is extensively driven by the association of diabetes with cardiovascular disease.

In conclusion, this study contributes further to understanding the drivers of poor outcomes in patients with diabetes admitted to hospital with COVID-19. Advancing age, multimorbidity (as crystallized in the CFS) and lower platelet count are important predictors of poor outcome in patients with COVID-19 admitted to hospital. The protective benefits of usage of antiplatelet agents are lost in patients with diabetes. There is no clear evidence that dysglycemia drives the increased risk of death/ICU admission among patients with diabetes, but rather the association of diabetes with other common medical conditions confers excess risk of poor outcomes.

#### Author affiliations

<sup>1</sup>Imperial Centre for Endocrinology, Imperial College Healthcare NHS Trust, London, UK

<sup>2</sup>Department of Metabolism, Digestion and Reproduction, Imperial College London, London, UK

<sup>3</sup>Imperial College Business School, Imperial College London, London, UK

<sup>4</sup>Department of Clinical Biochemistry, Imperial College Healthcare NHS Trust, London, UK

<sup>5</sup>Division of Medicine and Integrated Care, Imperial College Healthcare NHS Trust, London, UK

**Contributors** Conceptualization: CI-E, TMT, NEH and VS. Methodology: CI-E, AA, DCDH, TMT, NEH and VS. Investigation: CI-E, AA, WY, OI, JSK, RJS, EW, CS, NA, HB, NB, SB, ML, SP, ES, ET-B, AQ, MA-M, DCDH, NSO, VB, SM, TMT, NEH and VS. Formal analysis: WD and VS. Data curation: CI-E, AA, SM, TMT, NEH and VS. Writing-original draft preparation: CI-E and VS. Writing-review and editing: CI-E, WD, AA, WY, OI, JSK, RJS, EW, CS, NA, HB, NB, SB, ML, SP, ES, ET-B, AQ, MA-M, DCDH, NSO, VB, SM, TMT, NEH and VS. Visualization: CI-E, WD and VS. Guarantors: NEH and VS.

**Funding** The Section of Endocrinology and Investigative Medicine is funded by grants from the MRC, BBSRC and NIHR and is supported by the NIHR Biomedical Research Centre Funding Scheme. The views expressed are those of the authors and not necessarily those of the MRC, BBSRC, the NHS, the NIHR or the Department of Health and Social Care. CI-E is funded by an Imperial-BRC IPPRF Fellowship. JSK is funded by an NIHR Doctoral Research Fellowship (DRF-2017-10-042). SM is funded by a Future Leaders Mentorship Award from the European Association for the Study of Diabetes. TMT is funded by the MRC and NIHR. VS is funded by a Harry Keen Diabetes UK Fellowship.

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** This study was approved by the ICHNT Clinical Audit Office and was performed in accordance with the Declaration of Helsinki and UK Data Protection legislation.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. The data are deidentified patient data and requests for data should be made to v.salem@imperial.ac.uk. All data requests will be subject to relevant GDPR and ethics considerations.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

#### ORCID iDs

Chioma Izzi-Engbeaya <http://orcid.org/0000-0001-7599-0166>

Christine Shi <http://orcid.org/0000-0002-0507-0732>

Tricia M Tan <http://orcid.org/0000-0001-5873-3432>

Neil E Hill <http://orcid.org/0000-0002-4359-2646>

Victoria Salem <http://orcid.org/0000-0001-6463-7471>

#### REFERENCES

- Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* 2020;323:1239-1242.
- Guan W-J, Ni Z-Y, Hu Y, *et al*. Clinical characteristics of coronavirus disease 2019 in China. *N Engl J Med* 2020;382:1708-20.
- Yang X, Yu Y, Xu J, *et al*. Clinical course and outcomes of critically ill patients with SARS-CoV-2 pneumonia in Wuhan, China: a single-centered, retrospective, observational study. *Lancet Respir Med* 2020;8:475-81.
- Chen N, Zhou M, Dong X, *et al*. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020;395:507-13.
- Huang C, Wang Y, Li X, *et al*. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* 2020;395:497-506.
- Guan W-jie, Liang W-hua, Zhao Y, *et al*. Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J* 2020;55:2000547.
- Public Health England. *Diabetes prevalence and risk*, 2020.
- Booth CM, Matukas LM, Tomlinson GA, *et al*. Clinical features and short-term outcomes of 144 patients with SARS in the greater Toronto area. *JAMA* 2003;289:2801-9.
- Alqahtani FY, Aleanizy FS, Ali El Hadi Mohamed R, *et al*. Prevalence of comorbidities in cases of middle East respiratory syndrome coronavirus: a retrospective study. *Epidemiol Infect* 2018;1-5.
- Joensen LE, Madsen KP, Holm L, *et al*. Diabetes and COVID-19: psychosocial consequences of the COVID-19 pandemic in people with diabetes in Denmark-what characterizes people with high levels of COVID-19-related worries? *Diabet Med* 2020;37:1146-54.
- Zhu L, She Z-G, Cheng X, *et al*. Association of blood glucose control and outcomes in patients with COVID-19 and pre-existing type 2 diabetes. *Cell Metab* 2020;31:1068-77.
- Cariou B *et al*. Phenotypic characteristics and prognosis of inpatients with COVID-19 and diabetes: the CORONADO study. *Diabetologia* 2020.
- 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;8:823-33.

- 14 Rockwood K, Song X, MacKnight C, *et al*. A global clinical measure of fitness and frailty in elderly people. *CMAJ* 2005;173:489–95.
- 15 Fan J, Li R. Variable selection via nonconcave penalized likelihood and its oracle properties. *J Am Stat Assoc* 2001;96:1348–60.
- 16 Liu X, Song X, Xie S, *et al*. Variable selection for frailty transformation models with application to diabetic complications. *Canadian Journal of Statistics* 2016;44): :375–94.
- 17 Wang J-H, Pan C-H, Chang I-S, *et al*. Penalized full likelihood approach to variable selection for COX's regression model under nested case-control sampling. *Lifetime Data Anal* 2020;26:292–314.
- 18 NHS Digital. *National Diabetes Inpatient Audit (NaDIA) - 2017, 2018*.
- 19 Barron E *et al*. Type 1 and type 2 diabetes and COVID-19 related mortality in England: a whole population study 2020.
- 20 Wang S, Ma P, Zhang S, *et al*. Fasting blood glucose at admission is an independent predictor for 28-day mortality in patients with COVID-19 without previous diagnosis of diabetes: a multi-centre retrospective study. *Diabetologia* 2020;63:2102–11.
- 21 RECOVERY Collaborative Group, Horby P, Lim WS, *et al*. Dexamethasone in Hospitalized Patients with Covid-19 - Preliminary Report. *N Engl J Med* 2020. doi:
- 22 Zhou F, Yu T, Du R, *et al*. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* 2020;395:1054–62.
- 23 Carey IM, Critchley JA, DeWilde S, *et al*. Risk of infection in type 1 and type 2 diabetes compared with the general population: a matched cohort study. *Diabetes Care* 2018;41:513–21.
- 24 CDC COVID-19 Response Team. Preliminary Estimates of the Prevalence of Selected Underlying Health Conditions Among Patients with Coronavirus Disease 2019 - United States, February 12-March 28, 2020. *MMWR Morb Mortal Wkly Rep* 2020;69:382–6.
- 25 Chong E, Chan M, Tan HN, *et al*. COVID-19: use of the clinical frailty scale for critical care decisions. *J Am Geriatr Soc* 2020;68:E30–2.
- 26 Hewitt J, Carter B, Vilches-Moraga A, *et al*. The effect of frailty on survival in patients with COVID-19 (cope): a multicentre, European, observational cohort study. *Lancet Public Health* 2020;5:e444–51.
- 27 Tang N, Bai H, Chen X, *et al*. Anticoagulant treatment is associated with decreased mortality in severe coronavirus disease 2019 patients with coagulopathy. *J Thromb Haemost* 2020;18:1094–9.
- 28 Lippi G, Plebani M, Henry BM. Thrombocytopenia is associated with severe coronavirus disease 2019 (COVID-19) infections: a meta-analysis. *Clin Chim Acta* 2020;506:145–8.
- 29 Magro C, Mulvey JJ, Berlin D, *et al*. Complement associated microvascular injury and thrombosis in the pathogenesis of severe COVID-19 infection: a report of five cases. *Transl Res* 2020;220:1–13.
- 30 Middeldorp S *et al*. Incidence of venous thromboembolism in hospitalized patients with COVID-19. *J Thromb Haemost* 2020.
- 31 Poissy J *et al*. Pulmonary embolism in COVID-19 patients: awareness of an increased prevalence. *Circulation* 2020.
- 32 Thachil J, Juffermans NP, Ranucci M, *et al*. ISTH DIC Subcommittee communication on anticoagulation in COVID-19. *J Thromb Haemost* 2020;18:2138–44.
- 33 Williamson E *et al*. OpenSAFELY: factors associated with COVID-19 death in 17 million patients. *Nature* 2020.
- 34 Huang I, Lim MA, Pranata R. Diabetes mellitus is associated with increased mortality and severity of disease in COVID-19 pneumonia - A systematic review, meta-analysis, and meta-regression. *Diabetes Metab Syndr* 2020;14:395–403.
- 35 Palaiodimos L *et al*. Diabetes is associated with increased risk for in-hospital mortality in patients with COVID-19: a systematic review and meta-analysis comprising 18,506 patients 2020.

## 1 Supplementary Methods

### 2 STATISTICAL ANALYSES

3 Multiple logistic regression was chosen as the natural approach to model the  
4 probability of a binary outcome (death or ICU admission within 30 days of diagnosis).  
5 To neutralise the effects of extreme values all continuous variables were winsorized,  
6 at the 99% level in the case of the regression on all patients and at 97% in the case  
7 of the cohort of patients with diabetes (given their reduced sample size). We ran the  
8 first regression including all the potentially relevant predictors to avoid the problem of  
9 omitted variable bias. Estimated coefficients can be inconsistent if a relevant  
10 regressor is selected out of the model. Unlike in linear models, estimated coefficients  
11 in logistic regressions cannot be interpreted as marginal effects. The latter can be  
12 calculated alongside their standard errors. For example, the marginal effect of a  
13 generic regressor  $X_i$  measures the expected change on  $Pr$  (primary outcome  
14 measure) after a unit increase of  $X_i$ , holding all other  $X$ s at their sample averages.  
15 To enhance the interpretability and select the truly significant variables in our  
16 regression model, we have resorted to a regularised regression, which solves a  
17 general problem of the type:

$$\min_{\beta_0, \beta} \left( \frac{1}{2n} \sum_{i=1}^n (y_i - \beta_0 - x_i^\top \beta)^2 + \sum_{j=1}^p p_\lambda (|\beta_j|) \right)$$

18

19 where  $y_i$  is the response variable,  $x_i$  is the vector of  $p$  predictors,  $n$  is the number of  
20 observations and the  $\beta$ s are the parameters to estimate. The main difference  
21 between Lasso and the standard minimization involved in least squares regressions  
22 is the presence of the “penalty”  $\lambda \sum_{j=1}^p |\beta_j|$ , where  $\lambda$  is a non-negative tuning  
23 parameter. This penalty reduces the number of regressors and, therefore, serves the  
24 purpose of regularizing the problem and selecting the important predictors at the  
25 same time.

26 One of the first examples of regularized regression techniques is LASSO,  
27 popularized in the statistical literature by Tibshirani [1] and used in medical  
28 statistics.[2, 3] In many empirically relevant scenarios, the LASSO has been shown  
29 to perform better than stepwise selection methods.[4] In this paper we have resorted

30 to the SCAD penalty proposed by Fan et al.,[5] which has two important advantages:  
31 1. It doesn't penalize large coefficients; 2. It allows at the same time variable  
32 selection, estimation and inference, hence avoiding the need to re-estimate the  
33 model after selecting the relevant important predictors and controlling the errors  
34 involved in variable selection, in contrast to stepwise methods.

35 The penalty function in this model is defined by:

36

$$p'_{\lambda}(|\beta|) = \lambda \left\{ I(|\beta| \leq \lambda) = \frac{(a\lambda - |\beta|)_+}{(a-1)\lambda} I(|\beta| > \lambda) \right\}$$

37

38

39 The choice of  $\lambda$  is crucial for the performance of the methodology. As  $\lambda$  increases,  
40 the number of nonzero components in the vector  $\beta$  decreases. To avoid the danger  
41 of overfitting, we have used a tenfold cross validation technique to select  $\lambda$ .

42 Supplementary Tables

43 SUPPLEMENTARY TABLE S1 – Demographic and Clinical Features of 889 Patients  
 44 Treated for COVID-19 Patients in ICHNT Between March and April 2020 (Univariate  
 45 Analysis for primary outcome of death or Intensive Care Unit admission within 30  
 46 days of swab positive diagnosis).

<b>WHOLE COHORT</b>						
<i>Clinical feature</i>		<i>Number of patients with data</i>	<i>Prevalence (n)</i>	<i>Odds ratio of primary outcome (OR)</i>	<i>95% confidence interval (CI)</i>	<i>p-value</i>
Female		889	40% (355)	0.69	0.52 to 0.91	<b>0.01</b>
Age (years)		889				
	18-49		18% (164)	0.40	0.24 to 0.67	<b>&lt;0.0001</b>
	50-59		18% (157)	1	REFERENCE	
	60-69		17% (147)	1.11	0.74 to 1.84	0.55
	70-79		22% (193)	1.33	0.86 to 2.03	0.23
	>80		26% (228)	1.41	0.92 to 2.15	0.21
Weight (kg)		693				
	<60		19% (134)	1.83	1.48 to 2.26	<b>&lt;0.0001</b>
	60-90		58% (399)	1	REFERENCE	
	>90		23% (160)	1.02	0.88 to 1.09	0.77
Ethnicity		865				
	White		38% (327)	1	REFERENCE	
	Black		17% (146)	1.21	0.81 to 1.81	0.40
	South Asian		11% (97)	1.04	0.64 to 1.66	>0.9
	All non-White		62% (538)	1.35	1.01 to 1.81	<b>0.04</b>
Smoking status		484				
	Current/ex-smoker		189	1	REFERENCE	
	Never smoked		295	1.38	0.93 to 2.03	0.11
Co-morbidities		889		<i>OR (co-morbidity vs without co-morbidity)</i>		
	Hypertension		47% (418)	2.41	1.82 to 3.20	<b>&lt;0.001</b>
	Diabetes		38% (337)	1.68	1.27 to 2.21	<b>&lt;0.001</b>
	Hyperlipidaemia		33% (296)	1.56	1.16 to 2.08	<b>0.04</b>
	Ischaemic heart		16% (144)	2.25	1.58 to 3.23	<b>&lt;0.001</b>

	disease					
	Stroke		13% (120)	0.98	0.66 to 1.46	>0.9
	Heart failure		10% (87)	2.49	1.60 to 3.85	<b>&lt;0.001</b>
	COPD		9% (80)	1.12	0.71 to 1.81	0.63
	Active cancer		9% (79)	1.15	0.72 to 1.86	0.62
	Pre-COVID-19 eGFR <60mL/kg/min	694		1.40	1.25 to 1.56	<b>&lt;0.001</b>
Pre-admission medications		889		<i>OR (taking specified medication vs not taking the medication)</i>		
	Statin		42% (373)	1.49	1.13 to 1.97	<b>0.006</b>
	Anti-platelet drug		23% (204)		0.92 to 1.75	0.14
	ACE inhibitor		17% (151)	1.57	1.10 to 2.23	<b>0.013</b>
	Angiotensin II receptor blocker		13% (116)	1.43	0.95 to 2.12	0.09
CFS Score (1-9)		756				
	1 - 2		40% (300)	1	REFERENCE	
	3 - 4		32% (242)	1.50	1.05 to 2.15	<b>0.03</b>
	5 - 6		15% (110)	1.75	1.20 to 2.52	<b>0.003</b>
	≥7		13% (104)	2.34	1.46 to 3.75	<b>0.0004</b>
eGFR at time of COVID-19 diagnosis (mL/kg/min)		880				
	>90		27% (240)	1	REFERENCE	
	60 - 90		31% (269)	1.60	1.06 to 2.41	<b>0.02</b>
	45 - 59		11% (95)	2.35	1.40 to 3.91	<b>0.001</b>
	30 - 44		10% (93)	4.28	2.60 to 7.21	<b>&lt;0.0001</b>
	15 - 29		9% (78)	3.02	1.78 to 5.16	<b>&lt;0.0001</b>
	<15		12% (105)	4.19	2.56 to 6.78	<b>&lt;0.0001</b>
Glucose at time of COVID-19 diagnosis		652				
<i>WHOLE COHORT</i>	<10 mmol/L		78% (508)	1	REFERENCE	
	≥10 mmol/L		22% (144)	2.04	1.39 to 2.95	<b>0.0002</b>
		394				

PATIENTS WITHOUT DIABETES	<10 mmol/L		94% (369)	1	REFERENCE	
	≥10 mmol/L		6% (25)	3.56	1.60 to 8.62	<b>0.004*</b>

47 Supplementary Table S1: 889 patients with confirmed (swab positive) COVID-19  
 48 were analysed and univariate odds ratios (OR) against a defined comparator  
 49 (REFERENCE) are provided for our predefined primary outcome of death or ICU  
 50 admission within 30 days of diagnosis. ACE inhibitor = angiotensin converting  
 51 enzyme inhibitor, CBG = capillary blood glucose, CFS = Clinical Frailty Scale, COPD  
 52 = Chronic obstructive pulmonary disease, DPPIV inhibitor = dipeptidyl peptidase-IV  
 53 inhibitor, eGFR = estimated glomerular filtration rate, GLP-1RA = glucagon-like  
 54 peptide-1 receptor agonist, HbA1c = glycated haemoglobin, SGLT2 inhibitor =  
 55 sodium-glucose co-transporter-2 inhibitor. Significant p-values for odds ratios using  
 56 Fishers' exact test are shown in **bold**. \*Survives correction for age and gender.

57 SUPPLEMENTARY TABLE S2 – Demographic and Clinical Features of 337 Patients  
 58 with Diabetes Mellitus treated for COVID-19 in ICHNT Between March and April  
 59 2020 (Univariate Analysis for primary outcome of death or Intensive Care Unit  
 60 admission within 30 days of swab positive diagnosis).

PATIENTS WITH DIABETES						
Diabetes diagnosis		889				
	Patients without diabetes		62% (552)	1	REFERENCE	
	Patients with diabetes		38% (337)	1.68	1.27 to 2.12	<b>0.0003</b>
Type of diabetes		337				
	Type 2		96% (324)	1	REFERENCE	
	Type 1		4% (13)	0.37	0.11 to 1.27	0.16
Duration of diabetes		189				
	<10 years		47% (88)	1	REFERENCE	
	≥10 years		53% (101)	1.82	1.03 to 3.34	0.054
Diabetes medication at time of COVID-19 diagnosis		337		<i>OR (taking specified medication vs not taking the medication)</i>		
	Insulin		31% (108)	0.88	0.55 to 1.42	0.64
	GLP-1RA		1% (5)	0.52	0.09 to 2.60	0.66
	Metformin		49% (169)	1.14	0.74 to 1.76	0.58
	Sulphonylurea		22% (74)	0.85	0.50 to 1.45	0.59
	SGLT2 inhibitor		7% (24)	0.66	0.30 to 1.52	0.40
	DPPIV inhibitor		27% (93)	1.27	0.79 to 2.05	0.39
	Total number of diabetes medications					
	0		18% (57)	1	REFERENCE	
	1		41% (139)	0.91	0.50 to 1.67	0.75
	2		28% (96)	0.97	0.49 to 1.83	>0.9
	≥3		13% (45)	0.95	0.42 to 2.13	>0.9
HbA1c		240				
	<58 mmol/mol		47% (112)	1	REFERENCE	
	≥58 mmol/mol		53% (128)	1.44	0.85 to 2.39	0.19
Glucose at time of		323				

COVID-19 diagnosis						
	<10 mmol/L		56% (181)	1	REFERENCE	
	≥10 mmol/L		44% (142)	1.65	1.07 to 2.56	<b>0.03</b>
Average CBG in 1 <sup>st</sup> 72hr after COVID-19 diagnosis		306				
	<10 mmol/L		64% (197)	1	REFERENCE	
	≥10 mmol/L		36% (109)	1.93	1.22 to 3.11	<b>0.008</b>
Interaction of Diabetes with IHD		337				
	Without IHD		73% (246)	1	REFERENCE	
	With IHD		27% (91)	1.18	1.03 to 1.37	<b>0.02</b>
Interaction of Diabetes with pre-COVID-19 eGFR (mL/kg/min)		337				
	eGFR >80		24% (82)	1	REFERENCE	
	eGFR ≤79		76% (255)	1.53	1.03 to 2.29	<b>0.04</b>
Interaction of Diabetes with Hypertension		337				
	Without hypertension		29% (99)	1	REFERENCE	
	With hypertension		71% (238)	1.12	0.80 to 1.58	0.54

61 Supplementary Table S2: Out of the 889 patients studied in total (see Table S1), this  
62 table shows the results for the 337 patients within that group who also had a  
63 diagnosis of Diabetes Mellitus (DM). Patients with confirmed (swab positive)  
64 COVID-19 were analysed and univariate odds ratios (OR) against a defined  
65 comparator (REFERENCE) are provided for our predefined primary outcome of  
66 death or ICU admission within 30 days of diagnosis. ACE inhibitor = angiotensin  
67 converting enzyme inhibitor, CBG = capillary blood glucose, DPPIV inhibitor =  
68 dipeptidyl peptidase-IV inhibitor, eGFR = estimated glomerular filtration rate, IHD =  
69 ischaemic heart disease. Significant p-values for odds ratios using Fishers' exact  
70 test are shown in **bold**.

71 SUPPLEMENTARY TABLE S3 - Biochemical Measurements and Clinical  
 72 Observations for 889 patients admitted with COVID-19 according to whether they met  
 73 the primary outcome of death or Intensive Care Unit admission within 30 days of  
 74 swab positive diagnosis.

	Measurement at time of diagnosis	Primary outcome (died/ICU admission within 30 days)			Survived/no ICU admission			p-value (primary outcome vs survived/no ICU)
		n	mean	SD	n	mean	SD	
Arterial Blood Gas (ABG) measurement	FiO <sub>2</sub> recorded on ABG	260	48.70	26.68	403	29.81	18.2	<0.0001
	Lactate (mmol/L)	260	1.83	1.78	403	1.536	1.028	0.0077
	pCO <sub>2</sub> (kPa)	258	4.93	1.36	402	5.078	2.746	0.4618
	pO <sub>2</sub> (kPa)	255	8.81	5.97	397	8.208	5.596	0.1448
	Base excess (mmol/L)	256	-0.41	17.52	397	-0.37	4.953	0.1049
	pH	259	7.41	0.097	401	7.421	0.08202	0.222
	Bicarbonate (mmol/L)	256	23.28	5.69	400	24.05	4.415	0.0962
Laboratory results	WCC (x10 <sup>9</sup> /L)	320	8.61	4.09	561	7.95	4.635	0.0015
	Hb (g/L)	320	125.70	23.15	562	129.3	20.96	0.0264
	Platelet count (x10 <sup>9</sup> /L)	319	218.20	104.70	562	238.4	105.5	0.0024
	Neutrophils (x10 <sup>9</sup> /L)	318	7.42	7.22	562	6.346	8.549	<0.0001
	Lymphocytes (x10 <sup>9</sup> /L)	317	0.98	0.997	562	1.219	1.206	<0.0001
	D-dimers (ng/ml)	221	3895	5292	332	2445	3671	<0.0001
	Ferritin (ng/mL)	233	2122	5214	358	1313	1824	<0.0001
	Na (mmol/L)	320	138.30	6.231	560	137	7.424	0.043
	K (mmol/L)	309	4.33	0.75	534	4.213	0.5569	0.0783
	eGFR (mL/min)	319	51.14	29.67	560	65.66	27.47	<0.0001
	eGFR pre-COVID19 baseline (mL/min)	236	60.38	29.93	457	72.64	46.48	<0.0001
	ALT (U/L)	295	53.08	155.10	515	39.28	57.06	0.6581
	CRP (mg/L)	314	156.9	104.40	542	94.33	88.34	<0.0001
	CK (U/L)	206	641.9	1184	311	391.8	1179	<0.0001
	LDH (U/L)	186	602.5	626	291	398.8	230.2	<0.0001
BNP (pg/mL)	184	423.1	1451	267	192.2	762.7	<0.0001	
Procalcitonin (ng/mL)	33	20.5	69.04	49	0.538	1.527	<0.0001	
HbA1c - most recent (mmol/mol)	160	54.18	20.28	274	53.27	21.22	0.4003	
Glucose	CBG – at time of COVID19 diagnosis	269	10.04	5.54	416	8.312	5.398	<0.0001
	Average CBG in first 72hr of COVID-19 diagnosis	244	9.256	3.59	332	7.795	3.187	<0.0001
Clinical observations at	Temperature (°C)	311	37.38	1.19	558	37.26	1.016	0.0441

diagnosis	Respiratory Rate (Breaths per minute)	312	25.54	8.82	556	22.03	6.049	<b>&lt;0.0001</b>
	Heart Rate (beats per minute)	314	95.91	20.34	558	92.42	18.37	<b>0.0093</b>
	Systolic BP (mmHg)	316	134.8	27.49	560	131.2	23.74	<b>0.0452</b>
	Diastolic BP (mmHg)	316	73.53	16.06	559	75.67	13.81	<b>0.0321</b>
	NEWS at arrival	300	6	3.44	540	4.05	3.642	<b>&lt;0.0001</b>

75 Supplementary Table S3: Unadjusted comparisons are shown for patients who met  
76 the primary outcome (death or ICU admission within 30 days of diagnosis) against  
77 those who survived and did not have an ICU admission: n = number of patients for  
78 whom that data set is available. Note the total number of patients arriving at the  
79 primary outcome was 323. ABG = arterial blood gas, ALT = alanine  
80 aminotransferase, BNP = brain natriuretic peptide, BP = blood pressure, CBG =  
81 capillary blood glucose, CRP = C-reactive protein, CK = creatinine kinase, eGFR =  
82 estimated glomerular filtration rate, FiO2 = fraction of inspired oxygen, Hb =  
83 haemoglobin, HbA1c = glycated haemoglobin, K = potassium, LDH = lactate  
84 dehydrogenase, Na = sodium, NEWS = National Early Warning Score, WCC = white  
85 cell count.

86 SUPPLEMENTARY TABLE S4 - Biochemical Measurements and Clinical  
 87 Observations for 889 patients admitted with COVID-19 according to whether or not  
 88 they had a pre-existing diagnosis of Diabetes Mellitus

	Measurement at time of diagnosis	Patients without diabetes mellitus (no DM)			Patients with diabetes mellitus (DM)			p-value (no DM vs DM)
		n	mean	SD	n	mean	SD	
Arterial Blood Gas (ABG) measurement	FiO2 recorded on ABG	403	36.4	23.2	260	38.51	24.64	0.2867
	Lactate (mmol/L)	402	1.58	1.4	261	1.765	1.347	<b>0.0031</b>
	pCO2 (kPa)	401	4.96	1.28	259	5.117	3.321	0.8276
	pO2 (kPa)	397	8.27	5.45	255	8.71	6.183	0.5349
	Base excess (mmol/L)	397	0.0348	5.05	256	-2.048	6.781	<b>0.0003</b>
	pH	401	7.43	0.0813	259	7.401	0.09645	<b>0.001</b>
	Bicarbonate (mmol/L)	398	24.3	4.69	258	22.83	5.233	<b>0.0022</b>
Laboratory results	WCC (x10 <sup>9</sup> /L)	545	8.12	4.61	337	8.308	4.194	0.276
	Hb (g/L)	546	130	22.1	337	124.8	21.09	<b>0.0006</b>
	Platelet count (x10 <sup>9</sup> /L)	545	232	111	337	230.4	95.88	0.6622
	Neutrophils (x10 <sup>9</sup> /L)	545	6.28	4.18	336	7.472	11.98	0.0805
	Lymphocytes (x10 <sup>9</sup> /L)	545	1.17	1.34	335	1.079	0.7064	0.8785
	D-dimers (ng/ml)	336	2556	3714	218	3742	5305	<b>0.0096</b>
	Ferritin (ng/mL)	358	1553	2368	234	1751	4896	0.6944
	Na (mmol/L)	544	138	7.62	337	136.9	5.944	<b>0.0008</b>
	K (mmol/L)	524	4.18	0.603	320	4.378	0.6716	<b>&lt;0.0001</b>
	eGFR (mL/min)	544	67	26.4	336	49.64	30.14	<b>&lt;0.0001</b>
	eGFR pre-COVID19 baseline (mL/min)	411	75.4	47.2	283	58.38	30.34	<b>&lt;0.0001</b>
	ALT (U/L)	494	46.7	113	317	40.54	88.75	0.0657
	CRP (mg/L)	527	106	94	330	135.8	104.3	<b>&lt;0.0001</b>
	CK (U/L)	307	510	1320	211	464.5	960.5	0.6773
	LDH (U/L)	293	480	413	185	475.4	482.1	0.9935
	BNP (pg/mL)	279	201	754	173	422.4	1489	0.0688
Procalcitonin (ng/mL)	55	8.96	49.4	27	7.774	33.11	<b>0.0469</b>	
HbA1c - most recent (mmol/mol)	194	39.8	7.14	240	64.73	21.66	<b>&lt;0.0001</b>	
Glucose	CBG – at time of COVID19 diagnosis	375	6.97	2.07	311	11.43	7.145	<b>&lt;0.0001</b>
	Average CBG in first 72hr of COVID-19 diagnosis	273	6.9	1.77	303	9.776	3.969	<b>&lt;0.0001</b>
Clinical observations at diagnosis	Temperature (°C)	537	37.3	1.07	333	37.27	1.101	0.7922
	Respiratory Rate per Breaths	539	23.3	7.41	330	23.34	7.278	0.6356

	minute)							
	Heart Rate (beats per minute)	541	93	19.1	332	94.72	19.24	0.256
	Systolic BP (mmHg)	543	130	24.5	334	136.1	25.99	<b>0.0011</b>
	Diastolic BP (mmHg)	542	75.9	14.5	334	73.24	14.84	<b>0.0053</b>
	NEWS at arrival	523	4.63	3.95	321	4.938	3.234	0.0515

89 Supplementary Table S4: Unadjusted comparisons are shown for patients who had  
90 diabetes mellitus (n=337) and those who did not (n=552). n = number of patients for  
91 whom that data set is available. Note the total number of patients arriving at the  
92 primary outcome was 323. ABG = arterial blood gas, ALT = alanine  
93 aminotransferase, BNP = brain natriuretic peptide, BP = blood pressure, CBG =  
94 capillary blood glucose, CRP = C-reactive protein, CK = creatinine kinase, eGFR =  
95 estimated glomerular filtration rate, FiO2 = fraction of inspired oxygen, Hb =  
96 haemoglobin, HbA1c = glycated haemoglobin, K = potassium, LDH = lactate  
97 dehydrogenase, Na = sodium, NEWS = National Early Warning Score, WCC = white  
98 cell count.

99 SUPPLEMENTARY TABLE S5 - Biochemical Measurements and Clinical  
 100 Observations for 337 patients who had Diabetes Mellitus and were admitted with  
 101 COVID-19, according to whether they met the primary outcome of death or Intensive  
 102 Care Unit admission within 30 days of swab positive diagnosis.

	Measurement at time of diagnosis	Patients with Diabetes who suffered the primary outcome (died/ICU admission within 30 days) (n=149)			Patients with Diabetes who survived/no ICU admission (n=188)			p-value (primary outcome vs survived/no ICU)
		n	mean	SD	n	mean	SD	
Arterial Blood Gas (ABG) measurement	FiO <sub>2</sub> recorded on ABG	120	48.03	26.32	140	30.36	19.83	<b>&lt;0.0001</b>
	Lactate (mmol/L)	120	1.868	1.608	141	1.678	1.075	0.5536
	pCO <sub>2</sub> (kPa)	119	4.787	1.252	140	5.397	4.355	0.1177
	pO <sub>2</sub> (kPa)	117	10.04	7.493	138	7.583	4.533	<b>0.003</b>
	Base excess (mmol/L)	118	-2.86	7.946	138	-1.354	5.533	0.1905
	pH	119	7.399	0.09554	140	7.403	0.09751	0.7566
	Bicarbonate (mmol/L)	119	22.14	5.748	139	23.42	4.689	0.0853
Laboratory results	WCC (x10 <sup>9</sup> /L)	149	8.411	3.807	188	8.226	4.485	0.2624
	Hb (g/L)	149	123.4	22.92	188	126	19.51	0.2594
	Platelet count (x10 <sup>9</sup> /L)	149	211.4	84.43	188	245.4	101.8	<b>0.0011</b>
	Neutrophils (x10 <sup>9</sup> /L)	148	7.732	9.643	188	7.267	13.56	<b>0.008</b>
	Lymphocytes (x10 <sup>9</sup> /L)	147	0.9442	0.5931	188	1.184	0.7689	<b>0.0006</b>
	D-dimers (ng/ml)	108	4703	6244	110	2800	3995	<b>0.0298</b>
	Ferritin (ng/mL)	113	2230	6798	121	1304	1727	0.107
	Na (mmol/L)	149	137.7	6.015	188	136.3	5.834	<b>0.0346</b>
	K (mmol/L)	145	4.404	0.7085	175	4.357	0.6406	0.563
	eGFR (mL/min)	148	43.41	29.4	188	54.55	29.89	<b>0.0008</b>
	ALT (U/L)	142	50.71	127.6	175	32.29	30.81	0.1423
	CRP (mg/L)	147	162.9	102	183	114.1	101.3	<b>&lt;0.0001</b>
	CK (U/L)	102	716.6	1304	109	228.6	296.3	<b>0.0002</b>
	LDH (U/L)	87	584	658.6	98	379	190.7	<b>0.0036</b>
	BNP (pg/mL)	94	570.6	1944	80	248.4	592.5	<b>0.0483</b>
Procalcitonin (ng/mL)	16	12.68	42.85	11	0.64	1.096	<b>0.0299</b>	
HbA1c - most recent (mmol/mol)	99	63.56	20.14	141	65.55	22.71	0.968	
Glucose	CBG – at time of COVID19 diagnosis	140	12.28	6.537	171	10.74	7.555	<b>0.0021</b>
	Average CBG in first 72hr of COVID-19 diagnosis	141	10.5	3.937	162	9.141	3.898	<b>0.0004</b>
Clinical observations at diagnosis	Temperature (°C)	146	37.33	1.201	187	37.22	1.018	0.3678
	Respiratory Rate per minute (Breaths per minute)	146	25.34	8.627	184	21.76	5.53	<b>&lt;0.0001</b>

Heart Rate (beats per minute)	146	97.2	21.48	186	92.78	17.1	<b>0.0376</b>
Systolic BP (mmHg)	147	137.3	27.68	187	135.1	24.62	0.4514
Diastolic BP (mmHg)	147	72.5	15.78	187	73.82	14.07	0.4181
NEWS at arrival	142	6.07	3.37	183	4.06	2.837	<b>&lt;0.0001</b>

103 Supplementary Table S5: Unadjusted comparisons are shown for patients with  
 104 diabetes (total n=337) who met the primary outcome (death or ICU admission within  
 105 30 days of diagnosis, n=149) against those who survived and did not have an ICU  
 106 admission (n=188): n = number of patients for whom that data set is available. Note  
 107 the total number of patients arriving at the primary outcome was 323. ABG = arterial  
 108 blood gas, ALT = alanine aminotransferase, BNP = brain natriuretic peptide, BP =  
 109 blood pressure, CBG = capillary blood glucose, CRP = C-reactive protein, CK =  
 110 creatinine kinase, eGFR = estimated glomerular filtration rate, FiO2 = fraction of  
 111 inspired oxygen, Hb = haemoglobin, HbA1c = glycated haemoglobin, K = potassium,  
 112 LDH = lactate dehydrogenase, Na = sodium, NEWS = National Early Warning Score,  
 113 WCC = white cell count.

114 SUPPLEMENTARY TABLE S6 - Multivariate Regression Analysis with SCAD  
 115 sensitivity analysis to determine the important contributing regressors against the  
 116 Primary Outcome of Death/ICU Admission within 30 Days of COVID-19 Diagnosis  
 117 (n=719 patients)

<i>Regressor</i>	<i>Estimate</i>	<i>SE</i>	<i>p-value</i>	<i>Marginal Effect (%)</i>
Age	0.031	0.012	<b>0.009</b>	<b>0.5</b>
Gender	0.595	0.249	<b>0.017</b>	<b>10.3</b>
Ischaemic Heart Disease	0.002	0.017	0.893	0.04
Hypertension	0.147	0.331	0.657	2.6
Antiplatelet drug	-0.383	0.354	0.278	-6.7
Clinical Frailty Score	0.147	0.072	<b>0.043</b>	<b>2.5</b>
White cell count	0.035	0.049	0.472	0.6
Haemoglobin	-0.009	0.014	0.528	-0.2
Platelets	-0.004	0.002	<b>0.038</b>	<b>-0.1</b>
Neutrophils	0.000	0.015	0.977	<0.001
Serum sodium	0.044	0.025	0.077	0.8
Serum potassium	0.048	0.167	0.772	0.8
eGFR on diagnosis	-0.008	0.011	0.460	-0.1
C-reactive protein	0.002	0.002	0.298	0.0
Temperature at diagnosis	-0.052	0.276	0.851	-0.9
Respiratory rate on diagnosis	0.033	0.023	0.149	0.6
Heart Rate on diagnosis	0.006	0.008	0.427	0.1
Systolic Blood Pressure	0.011	0.006	<b>0.041</b>	<b>0.2</b>
Diastolic Blood Pressure	-0.004	0.010	0.657	-0.1
NEWS at diagnosis	-0.004	0.080	0.959	-0.1
Inspired Oxygen delivered at diagnosis	-0.013	0.007	0.087	-0.2
Oxygen saturations at diagnosis	-0.035	0.030	0.236	-0.6
Maximum inspired oxygen required	0.067	0.005	<b>&lt;0.0001</b>	<b>1.2</b>

118 Supplementary Table S6: Unselected multivariate logistic (Logit) analysis of all  
 119 variables that were collected for patients admitted with swab positive COVID 19 with  
 120 the primary outcome of death or ICU admission within 30 days (719 patients) – that  
 121 data is shown in Table 1. Thereafter a sensitivity analysis was performed to select  
 122 the truly important variables in our regression. We used a Smoothly Clipped  
 123 Absolute Deviation (SCAD) regularised regression technique which delivered 23  
 124 variables independently selected to be the most important drivers of the model, and  
 125 new p-values are recreated to adjust for this analysis. Refer to Table 1 for details on  
 126 the other statistical measures presented. eGFR = estimated glomerular filtration rate,  
 127 NEWS = National Early Warning Score.

128 SUPPLEMENTARY TABLE S7 - Multivariate Regression Analysis with SCAD  
 129 sensitivity analysis to determine the important contributing regressors against the  
 130 Primary Outcome of Death/ICU Admission within 30 Days of COVID-19 Diagnosis in  
 131 Patients with Diabetes Mellitus (n=268 patients)

<i>Regressor</i>	<i>Estimate</i>	<i>SE</i>	<i>p-value</i>	<i>Marginal Effect (%)</i>
Age	0.041	0.017	<b>0.016</b>	<b>0.9</b>
Gender	0.024	0.105	0.822	0.5
White cell count	-0.001	0.027	0.970	0.0
Haemoglobin	-0.012	0.011	0.265	-0.3
Platelets	-0.007	0.003	<b>0.008</b>	<b>-0.2</b>
Neutrophils	0.000	0.006	0.971	0.0
Serum sodium	0.000	0.008	0.982	0.0
eGFR on diagnosis	-0.008	0.007	0.236	-0.2
C-reactive protein	0.000	0.002	0.836	0.0
Capillary blood glucose on diagnosis	0.024	0.030	0.422	0.6
Temperature on diagnosis	-0.017	0.077	0.824	-0.4
Respiratory rate on diagnosis	0.040	0.031	0.196	0.9
Heart rate on diagnosis	-0.004	0.010	0.682	-0.1
Systolic blood pressure	0.004	0.011	0.713	0.1
Diastolic blood pressure	0.006	0.022	0.800	0.1
Inspired oxygen delivered at diagnosis	0.003	0.012	0.773	0.1
Oxygen saturations at diagnosis	-0.005	0.047	0.908	-0.1
Maximum inspired oxygen required	0.057	0.008	<b>&lt;0.0001</b>	<b>1.3</b>

132 Supplementary Table S7: Unselected multivariate logistic (Logit) analysis of all  
 133 variables that were collected for patients admitted with swab positive COVID 19 who  
 134 had diabetes mellitus, as applied to the primary outcome of death or ICU admission  
 135 within 30 days (268 patients) – that data is shown in Table 2. Thereafter a sensitivity  
 136 analysis was performed to select the truly important variables in our regression. We  
 137 used a Smoothly Clipped Absolute Deviation (SCAD) regularised regression  
 138 technique which delivered 19 variables independently selected to be the most  
 139 important drivers of the model, and new p-values are recreated to adjust for this  
 140 analysis. Refer to Table 1 for details on the other statistical measures presented.  
 141 eGFR = estimated glomerular filtration rate.

142 Supplementary References

- 143 1. Tibshirani, R., *Regression Shrinkage and Selection via the Lasso*. Journal of the Royal  
144 Statistical Society. Series B (Methodological), 1996. **58**(1): p. 267-288.
- 145 2. Tibshirani, R., *The lasso method for variable selection in the Cox model*. Stat Med, 1997.  
146 **16**(4): p. 385-95.
- 147 3. Yengo, L., et al., *Impact of statistical models on the prediction of type 2 diabetes using non-*  
148 *targeted metabolomics profiling*. Mol Metab, 2016. **5**(10): p. 918-925.
- 149 4. Hastie, T., R. Tibshirani, and R. Tibshirani, *Best Subset, Forward Stepwise, or Lasso? Analysis*  
150 *and Recommendations Based on Extensive Comparisons*. Preprint at  
151 <https://arxiv.org/abs/1707.08692>, 2020.
- 152 5. Fan, J. and R. Li, *Variable selection via nonconcave penalized likelihood and its oracle*  
153 *properties*. Journal of the American Statistical Association, 2001. **96**(456): p. 1348-1360.