



## OPEN The relationship between hyperglycaemia on admission and patient outcome is modified by hyperlactatemia and diabetic status: a retrospective analysis of the eICU collaborative research database

Oisín Fitzgerald<sup>1</sup>✉, Oscar Perez-Concha<sup>1</sup>, Blanca Gallego-Luxan<sup>1</sup>, Lachlan Rudd<sup>2</sup> & Louisa Jorm<sup>1</sup>

Both blood glucose and lactate are well-known predictors of organ dysfunction and mortality in critically ill patients. Previous research has shown that concurrent adjustment for glucose and lactate modifies the relationship between these variables and patient outcomes, including blunting of the association between blood glucose and patient outcome. We aim to investigate the relationship between ICU admission blood glucose and hospital mortality while accounting for lactate and diabetic status. Across 43,250 ICU admissions, weighted to account for missing data, we assessed the predictive ability of several logistic regression and generalised additive models that included blood glucose, blood lactate and diabetic status. We found that inclusion of blood glucose marginally improved predictive performance in all patients: AUC-ROC 0.665 versus 0.659 ( $p = 0.005$ ), with a greater degree of improvement seen in non-diabetics: AUC-ROC 0.675 versus 0.663 ( $p < 0.001$ ). Inspection of the estimated risk profiles revealed the standard U-shaped risk profile for blood glucose was only present in non-diabetic patients after controlling for blood lactate levels. Future research should aim to utilise observational data to estimate whether interventions such as insulin further modify this effect, with the goal of informing future RCTs of interventions targeting glycaemic control in the ICU.

Both blood glucose and lactate are well-known predictors of organ dysfunction and mortality in critically ill patients<sup>1,2</sup>. Glucose shows a U-shaped relationship with mortality with both hypo- and hyperglycaemia associated with poor outcomes<sup>3,4</sup>. The association between hospital mortality and lactate is particularly strong, with previous research demonstrating lactate has comparable predictive ability to the APACHE-II, SOFA and qSOFA scores in certain ICU populations<sup>5</sup>. Further, it has been shown in previous research that concurrent adjustment for glucose and lactate modifies the relationship between these variables and patient outcomes, including blunting of the association between blood glucose and patient outcome<sup>5-8</sup>.

The metabolic pathways of glucose and lactate are highly inter-connected. Lactate is an end-product of glycolysis, the oxygen independent cellular respiration of glucose, and is a major substrate for gluconeogenesis in the liver and kidney<sup>9</sup>. In previous work, our team has found that blood lactate is the most important non-glucose blood gas/laboratory factor in predicting future blood glucose in critically ill patients<sup>10</sup>. The mechanisms that lead to hyperglycaemia (> 180 mg/dL) and hyperlactatemia (> 2 mmol/L) in critical illness are complex and varied, with clinical interventions playing a role<sup>11-13</sup>. While tissue hypoxia leading to increased anaerobic

<sup>1</sup>Centre for Big Data Research in Health, Level 2, AGSM Building, UNSW Sydney, Kensington, NSW 2052, Australia. <sup>2</sup>Data and Analytics, eHealth NSW, 1 Reserve Road, St Leonards, NSW 2065, Australia. ✉email: o.fitzgerald@unsw.edu.au

respiration has commonly been seen as the primary cause of increased lactate levels, this view has been challenged, with both hyperlactatemia and hyperglycaemia also linked to metabolic alterations associated with immune activation<sup>9,14–16</sup>. A major question for clinicians and researchers is when components of this response are maladaptive requiring treatment.

The degree to which blood glucose (at varying thresholds) is a marker or mediator of poor patient outcomes has been debated<sup>17,18</sup>, with observational studies demonstrating lack of, or attenuated, association between blood glucose and patient outcome in certain circumstances<sup>5–8</sup>, and randomised control trials (RCTs) demonstrating that tight glycaemic control (a target 80–110 mg/dL) is associated with poorer outcomes than a less stringent target of < 180 mg/dL<sup>19</sup>. The pathway from short term glucose toxicity to poor outcomes in the critically ill has not been fully elucidated, with organ dysfunction as a result of glucose induced inflammation and oxidative damage one hypothesised route<sup>20,21</sup>. However, elevated lactate may also have deleterious effects, for instance through immunosuppression<sup>16,22</sup>, raising the potential that elevated blood glucose levels may be a marker of this effect.

Accordingly, we aim to investigate whether blood glucose is an independent predictor of hospital outcome while controlling for blood lactate across subgroups defined by diabetic status. To account for non-linear relationships, we use flexible semi-parametric statistical models which retain the benefit of being readily interpretable while accounting for non-linear effects<sup>23</sup>.

## Methods

**Patients and data sources.** Data for this study were sourced from the eICU collaborative research database (eICU-CRD) open access critical care database, de-identified to conform with the Health Insurance Portability and Accountability Act (HIPAA). eICU-CRD is a large multi-center critical care database holding data associated with 200,859 ICU stays admitted at 208 hospitals across the United States between 2014 and 2015<sup>24</sup>. As described in previous research<sup>25,26</sup> data from the eICU-CRD are generally of high quality, with common vital signs (such as blood glucose) and patient outcomes well recorded, but missing data creating challenges identifying which patients received complex interventions such as intubation, ventilation, and dialysis beyond the first day of ICU stay and in the calculation of risk scores that depend on non-bedside measurements.

**Inclusion criteria and data quality assessment.** We restrict our analysis to non-elective adult patients (over 18-years-old) who had an ICU stay more than 12 h in duration, received an APACHE-IV score and were not admitted for diabetic ketoacidosis or hyperosmolar hyperglycaemic state. From this cohort we selected all patients who had at least one glucose and lactate measurement within –12 to +24 h of their ICU admission (see Fig. 1).

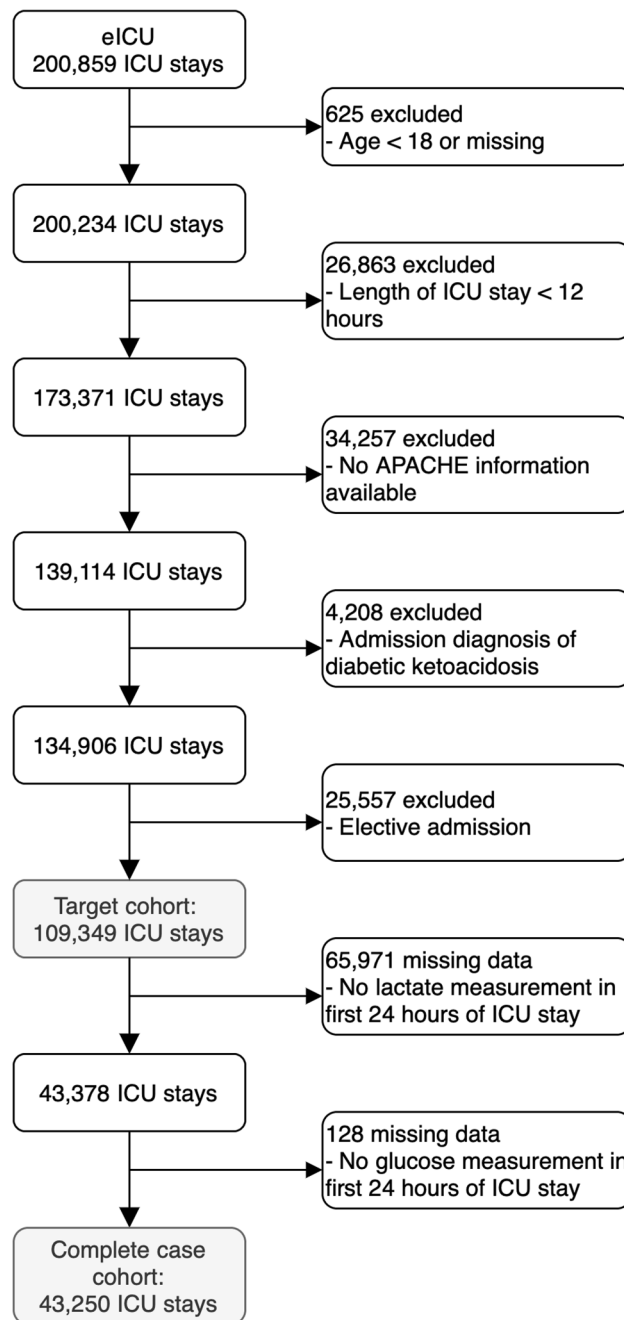
**Data extraction.** All data extraction and transformation processes were carried out using the R packages *bigquery*<sup>27</sup> and *data.table*<sup>28</sup>. All queries are freely available at the project code repository (linked below). The following variables on admission (–12 to +24 h of index admission time) were extracted from the database: ICU type, age, gender, ethnicity, weight, BMI, diabetic status, Elixhauser comorbidity index, APACHE admission diagnosis, APACHE-IV score, lactate, glucose, bilirubin, potassium, sodium, chloride, creatinine, blood urea nitrogen, calcium, along with indication of the use or prescription of insulin, intubation and ventilation. Extracted outcomes included in-hospital and in-ICU mortality and length of ICU stay.

**Ethics statement and reporting standards.** Use of the eICU-CRD dataset is exempt from institutional review board approval due to the retrospective design, lack of direct patient intervention, and the security schema, for which the re-identification risk was certified as meeting safe harbor standards by an independent privacy expert (Privacert, Cambridge, MA) (Health Insurance Portability and Accountability Act Certification no. 1031219-2). This includes a waiver of requirement for patient consent. Negligible risk ethics approval to ensure the analysis protocols and methods met relevant guidelines and regulations was obtained from UNSW Sydney—HC220829. We used the STROBE reporting guidelines<sup>29</sup> to check we provided an accurate and complete report of the study.

**Statistical analysis.** *Matching blood glucose and lactate measurements.* As blood lactate is measured less frequently than blood glucose, we construct our blood glucose and lactate variables by: (1) selecting the blood lactate measurement closest to ICU admission and (2) taking the average of all blood glucose measurements within 1–12 h of this blood lactate measurement. The size of the blood glucose averaging window is the minimum value (at 1 h increments) from 1 to 12 such that at least one blood glucose measurement is available.

*Descriptive statistics.* We give a descriptive analysis of the data using graphs of the key variables and through summarising the dataset in a table. Graphically we report the univariate and bivariate relationship between blood glucose and blood lactate, stratified by diabetic status and hospital mortality status. To avoid overplotting we show the bivariate relationships using contour plots for the stratified results. In the tables, continuous variables are reported using the mean and standard deviation. Categorical variables are reported as percentages. Data are reported stratified by whether a lactate measurement was available.

*Statistical modelling of missingness.* We use inverse probability weighting to account for potential selection biases introduced by the irregular measurement of blood lactate<sup>30</sup>. This approach uses the complete cases to create a pseudo-population designed to mimic the characteristics of the target population (in this case adults with an ICU stay over 12 h and an APACHE-IVa score<sup>31</sup>). We use a logistic regression model to calculate the



**Figure 1.** Impact of the inclusion criteria and missing data on extraction of the target and final complete case cohort from the eICU-CRD database.

probability that blood lactate would be measured in our target population using the following covariates: age, APACHE-IVa score, operative admission, an APACHE admission diagnosis of sepsis, APACHE admission organ system and ICU type. We assess the model fit using the area under the ROC curve (AUC-ROC) and a calibration plot. For each row in the complete case dataset the weights are calculate as the inverse of the model prediction.

We performed several sensitivity analyses to ensure our results are not overly impacted by our approach to account for missing data. We use a machine learning model (XGBoost<sup>32</sup>) to additionally calculate the missingness weights using the same set of variables as the logistic regression model, along with all non-lactate laboratory results (see Data Extraction above). To avoid overfitting bias this was done using cross-validated out of sample prediction<sup>33</sup>. Additionally, we develop a linear generalised additive model (GAM) and XGBoost imputation models, that were fit on the subsample of observations with a lactate measurement. These use the same variables sets and methodologies (i.e., out of sample prediction for the XGBoost model) as the missingness models, and in the imputation analysis their prediction replace the unmeasured lactate values for the 65,971 observations with no recorded blood lactate.

**Statistical modelling of hospital mortality.** The aim of the statistical modelling of hospital mortality was: (1) to assess the degree to which blood glucose is predictive of mortality when controlling for blood lactate and diabetic status (model selection step) and (2) describe the functional form of this relationship (model interpretation step).

**Models.** We use logistic and GAMs to build the models. GAMs are additive models that allow the predictor effects to vary non-linearly with the level of the covariate by transforming the variable through some smooth function  $f$ , e.g., a spline parameterisation. For a given basis dimension, GAMs incorporate the penalization of the spline and model estimation into a single process reducing the need for model hyperparameter tuning compared to pre model fitting spline transformation of a regression variable<sup>23</sup>. We use thin-plate regression splines<sup>34</sup>. In cases where the smooth spline is a two-dimensional interaction effect we denote this by (*var1:var2*). For the logistic regression (LR) models, we transformed glucose and lactate into categorical (binned) variables using the cut points from Freire Jorge, Wieringa<sup>8</sup>. We fit the following models, for all observations and stratified by diabetic status, using the iterative fitting method introduced in Wood<sup>23</sup> for GAMs and maximum likelihood for logistic regression models:

1. LR: glucose
2. LR: lactate
3. LR: glucose + lactate
4. LR: glucose + lactate + (glucose : lactate)
5. GAM: glucose
6. GAM: lactate
7. GAM: glucose + lactate
8. GAM: (glucose : lactate)

**Model selection.** Mirroring the aims, the statistical analysis was performed in two steps: model selection and model interpretation. During the model selection step a weighted tenfold cross-validation (CV) procedure was used to assess model fit<sup>35</sup>. The splits were made using the patient health care system ID, ensuring that all ICU stays from a patient were either in the training or test dataset. The performance of the models was assessed using the missing data model weights to calculate weighted versions of accuracy, the Brier score (mean squared error), logarithmic loss, the AUC-ROC, and the area under the precision-recall curve (AUC-PR). The results are presented overall and stratified by diabetic status. We repeated this procedure (unweighted for the imputation models) for each of the missingness sensitivity analysis models described above (Statistical Modelling of Missingness), along with an additional weighted approach restricted to only patients with an admission diagnosis of sepsis.

**Model interpretation.** During the final step we refit the models on the entire dataset using the missingness weights, graphing relevant GAM non-linear effects for interpretation and comparison so LR coefficients. Additionally, as a sensitivity analysis we fit a XGBoost model using blood glucose, blood lactate and diabetic status and assess the resulting non-parametric estimate of the relationship between blood glucose and hospital mortality for various blood lactate levels by diabetic status.

**Code availability.** All analysis were performed using R version 4.2.1<sup>36</sup> using ggplot2<sup>37</sup> and gratia<sup>38</sup> for graphics, mgcv<sup>23</sup> and XGBoost<sup>32</sup> for statistical modelling, along with the packages listed previously. The project code is available at [www.github.com/oizin/lactate](https://www.github.com/oizin/lactate).

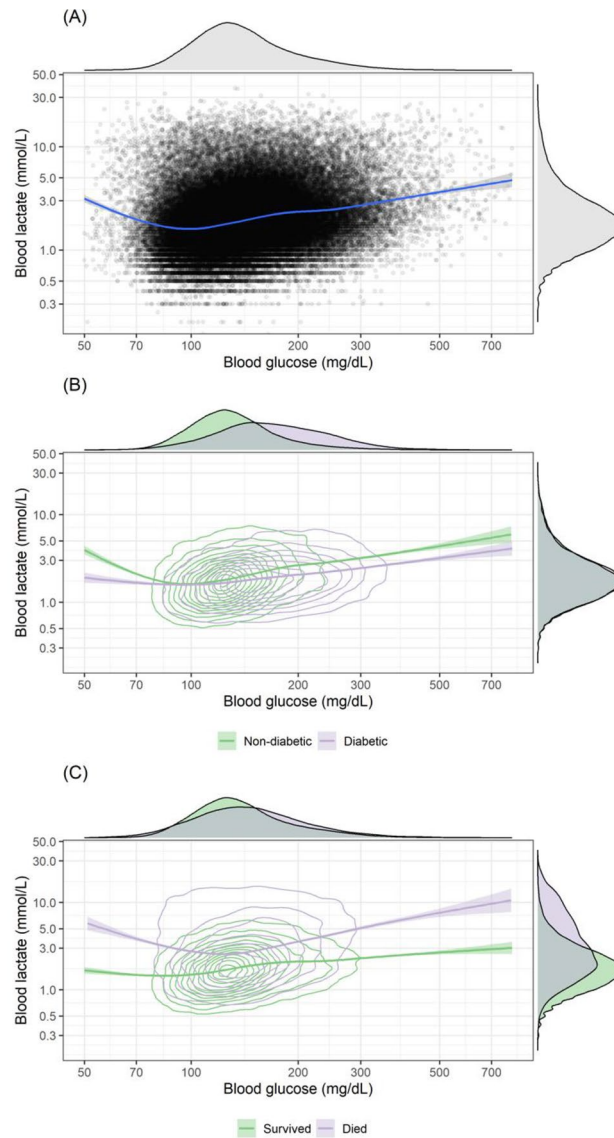
## Results

**Patient characteristics.** The characteristics of patients at ICU admission, over the first 24 h of their ICU stay and at discharge are shown in Table 1. The table enables comparison of those patients who had a lactate measurement taken with those who did not—with more analysis of the impact of missingness found in Appendix A (Fig. A1). Those who had a lactate measurement taken were on average more severely ill (Apache-IVa score: 68 vs 50), more likely to have an admission diagnosis of sepsis (33% vs 6%), more likely to be ventilated or receiving insulin (45% vs 24% and 38% vs 25%) and had poorer outcomes (hospital mortality of 16.2% vs 6.5%). The missing data weighting overcame these differences to a large degree, albeit with some residual differences. Comparing the weighted cohort to the overall cohort we see that the above differences in illness severity (Apache-IV score: 58 vs 57), sepsis diagnosis (17% vs 17%), ventilated (35% vs 32%), insulin (36% vs 30%), and outcomes (hospital mortality of 11.5% vs 10.3%) are greatly reduced. Our weighted analysis cohort is thus a marginally more severely ill version of our target cohort (Fig. 1). The mean lactate measurement of 2.3 mmol/L in the weighted cohort compares to values of 3.8 mmol/L (hospital mortality: 20.1%)<sup>5</sup>, 1.5 mmol/L (hospital mortality: 11.1%)<sup>8</sup> and 1.4 mmol/L (hospital mortality: 13.3%)<sup>6</sup> in previous literature. Additionally, we note that the hospital mortality rate in the 34,257 ICU stays without APACHE information was 8.7%, comparable with the target cohort with elective admissions included (hospital mortality percentage of 9.0%) and suggestive that this information could be considered missing at random.

Figure 2 outlines the univariate and bivariate relationship between blood glucose and blood lactate. As seen in Fig. 2A both blood glucose and blood lactate have right skewed distributions, even after log-transformation. There is a suggestion of a U-shaped relationship between blood glucose and blood lactate, with both hyperglycaemia and hypoglycaemia associated with hyperlactatemia. Stratifying by diabetic status the results are in line

	Overall	Weighted	Lactate measurement	
			Available	Not available
Patients (N)	104,867	42,540	42,540	64,057
Hospitals	189	189	189	189
Patient characteristics at ICU admission				
ICU stays (N)	109,349	43,378	43,378	65,971
Age (years) (mean (SD))	64(17)	64 (17)	64 (17)	63 (17)
Gender: female (%)	54	53	53	54
BMI (kg/m <sup>2</sup> ) (mean (SD))	28.7 (7.7)	28.7 (7.8)	28.7 (8.0)	28.9 (7.6)
Diabetic (%)	21.9	22	22.9	21.3
Ethnicity (%)				
Caucasian	76.7	77.2	77.3	76.2
African American	11.8	12.4	12.1	11.6
Hispanic	4	3.3	3.3	4.4
Asian	1.4	1.4	1.4	1.4
Other/unknown	5.5	5	5	5.9
ICU type (%)				
Cardiac ICU	7.8	7.3	7.3	8.2
CCU-CTICU	8.3	7.9	6.4	9.6
CSICU	2	2	2	2
CTICU	1.9	1.9	1.4	2.2
Med-Surg ICU	56.7	57	61.3	53.7
MICU	10.3	11	12.2	9.1
Neuro ICU	7.5	7.2	3.9	9.8
SICU	5.4	5.7	5.5	5.4
ICU admission type				
Operative	2.8	2.8	2.8	2.9
Admission diagnosis of sepsis (%)	16.9	17.3	32.8	6.4
Initial 24 h of ICU stay				
Glasgow coma score (mean (SD))	12.8 (3.7)	12.6 (3.8)	11.7 (4.3)	13.4 (3.2)
APACHE-IVa score (mean (SD))	57 (26)	58 (26)	68 (28)	50 (21)
Glucose measures in first 24 h				
Initial glucose (mg/dL) (mean (SD))	145 (74)	148 (78)	153 (84)	140 (65)
Matched glucose (mg/dL) (mean (SD))	143 (57)	147 (61)	151 (66)	137 (48)
Mean glucose (mg/dL) (mean (SD))	141 (51)	143 (51)	146 (54)	137 (48)
Maximum glucose (mg/dL) (mean (SD))	182 (98)	191 (103)	199 (111)	170 (87)
Glucose < 80 mg/dL (%)	15	17	19	12
Glucose < 50 mg/dL (%)	3	3	4	2
No glucose measurement in first 24 h	2099	128	128	1971
Lactate measures in first 24 h				
Initial lactate (mmol/L) (mean (SD))	n/a	2.3 (2.2)	2.6 (2.6)	n/a
Matched lactate (mmol/L) (mean (SD))	n/a	2.5 (2.4)	2.8 (2.7)	n/a
Mean lactate (mmol/L) (mean (SD))	n/a	2.2 (2.0)	2.5 (2.3)	n/a
Maximum lactate (mmol/L) (mean (SD))	n/a	2.7 (2.7)	3.1 (3.2)	n/a
Other lab measures in first 24 h				
Mean total bilirubin (mg/dL) (mean (SD))	1.1 (2.2)	1.1 (2.1)	1.2 (2.5)	0.9 (1.9)
Mean creatinine (mg/dL) (mean (SD))	1.6 (1.8)	1.7 (1.8)	1.9 (1.9)	1.5 (1.7)
Interventions in first 24 h				
Treated with insulin (%)	30.3	35.8	37.7	25.4
Ventilated (%)	32	35	45	24
Intubated (%)	24	25	36	16
Patient outcomes				
Hospital mortality (%)	10.3	11.5	16.2	6.5
ICU mortality (%)	6.3	7.1	10.7	3.4
Length of ICU stay (hours) (mean (SD))	79 (104)	84 (104)	95 (118)	68 (92)

**Table 1.** Descriptive statistics for all eligible patients (overall), the missing data weighted cohort and the overall cohort stratified by lactate availability.



**Figure 2.** The univariate and bivariate distributions of admission blood glucose and blood lactate with fitted loess curve for: (A) all data points; (B) stratified by diabetic status; and (C) stratified by hospital survival status.

with expectation (Fig. 2B). Patients with diabetes have an upward shift in their blood glucose distribution. In comparison, grouping by hospital outcome reveals those who died have a greater spread in blood glucose values, in line with evidence that both hypoglycaemia and hyperglycaemia are markers of mortality risk. In contrast, the relationship for blood lactate appears simpler, with a clear upward shift in the distribution of blood lactate for those who died.

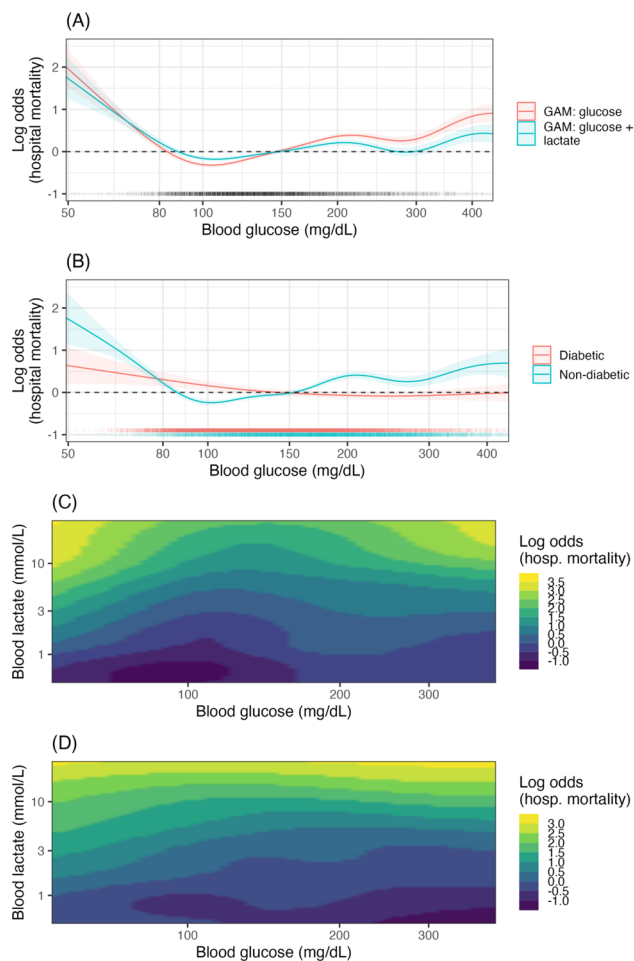
**Missingness model.** The coefficients, and measures of their uncertainty, for the variables we used to model missing lactate are shown in Appendix A (Table A2). As the model outcome is 1 if lactate is missing and 0 otherwise, a positive coefficient indicates increased likelihood of missingness when the variable's value increases. The results are in line with Table 1, with each unit reduction in the Apache-IVa score associated with a ~3% increase in the odds of missing lactate. Neuro ICU and operative patients are less likely to have a lactate measurement while sepsis patients are more likely. The model has good discrimination with an AUC-ROC of 0.788, illustrated in Appendix A Fig. A2A. Graphical analysis of model calibration (Fig. A2B) reveals that calibration is generally good, with some evidence that the lower probabilities are marginal over-estimates of the chance of missingness (Fig. A2A). As seen in Fig. A2A the distribution of the missingness probabilities is bimodal, with a broadly sepsis group clustered around 0.25 and non-sepsis around 0.75. The resulting weights are largely less than 10, with 75% less than 4 and 50% less than 2 (Fig. A2D). Results for the alternative XGBoost model and blood lactate imputation models can be found in Appendix A (Tables A3–A5).

**Model selection.** The results of the model cross-validation are given in Table 2 and show variation in performance across the models compared to the reference model *GAM: lactate*. Overall, the addition of blood glucose into the predictive models improve performance over lactate alone, with the GAMs that include blood glucose outperforming the reference (lactate only) model, with a paired *t-test* indicating a significant difference ( $p=0.005$ ) between the predictive performance of *GAM: lactate* (mean AUC-ROC: 0.659) and *GAM: glucose + lactate* (mean AUC-ROC: 0.664). This effect was most clear in non-diabetics, with a paired *t-test* indicating a significant difference ( $p < 0.001$ ) between the predictive performance of *GAM: lactate* (mean AUC-ROC: 0.663) and *GAM: glucose + lactate* (mean AUC-ROC: 0.675).. Further, the models illustrate that binning of covariates results in an inferior performance compared to the comparable GAM. As shown in Appendix A (Fig. A3–A7) these results were consistent across sensitivity analyses.

**Model interpretation.** Based on the results of the cross-validation we refit the GAM models on the data due to their similar performance. The regression spline effects are shown in Fig. 3, with the effect of blood glucose on mortality varying depending on which factors are adjusted for in the model. Adjustment for lactate alone results in a moderate attenuation of the association between hyperglycaemia (although not hypoglycaemia) and hospital outcome (Fig. 3A). Further adjustment for diabetic status reveals different risk profiles for diabetics and non-diabetics (Fig. 3B), with the effect of hyperglycaemia much reduced for diabetics. There is evidence of an interaction effect between glucose and lactate in non-diabetics with both hypo- and hyper-glycaemia associated with a poorer outcome compared to normo-glycaemia for a given blood lactate level (Fig. 3C). This interaction effect is attenuated for diabetics (Fig. 3D). These results are largely consistent with the XGBoost based sensitivity analysis reported in Appendix A (Fig. A9) which suggests an interaction effect whereby the risk profile for blood glucose is less attenuated at higher blood lactate values.

Model	AUC-ROC	AUC-PR	Log loss	Brier
Overall (N = 43,250)				
LR: glucose	0.558	0.137	0.355	0.102
LR: lactate	0.631	0.164	0.346	0.100
LR: glucose + lactate	0.638	0.171	0.345	0.099
LR: glucose + lactate + (glucose: lactate)	0.640	0.173	0.345	0.099
GAM: glucose	0.577	0.153	0.354	0.101
GAM: lactate	0.659	0.232	0.338	<b>0.097</b>
GAM: glucose + lactate	0.664	0.239	0.337	<b>0.097</b>
GAM: (glucose : lactate)	<b>0.665</b>	<b>0.244</b>	<b>0.336</b>	<b>0.097</b>
APACHE-IVa	0.823	0.412	0.263	0.076
Non-diabetic (N = 33,310)				
LR: glucose	0.577	0.153	0.357	0.103
LR: lactate	0.636	0.169	0.349	0.101
LR: glucose + lactate	0.650	0.184	0.347	0.101
LR: glucose + lactate + (glucose : lactate)	0.651	0.185	0.347	0.101
GAM: glucose	0.598	0.173	0.355	0.102
GAM: lactate	0.663	0.235	0.342	0.099
GAM: glucose + lactate	<b>0.675</b>	<b>0.252</b>	<b>0.339</b>	<b>0.098</b>
GAM: (glucose : lactate)	0.673	0.249	<b>0.339</b>	<b>0.098</b>
APACHE-IVa	0.803	0.369	0.265	0.075
Diabetic (N = 9940)				
LR: glucose	0.503	0.106	0.342	0.096
LR: lactate	0.614	0.147	0.334	0.095
LR: glucose + lactate	0.620	0.153	0.334	0.094
LR: glucose + lactate + (glucose: lactate)	0.614	0.151	0.334	0.095
GAM: glucose	0.529	0.126	0.341	0.096
GAM: lactate	<b>0.644</b>	<b>0.231</b>	<b>0.325</b>	<b>0.092</b>
GAM: glucose + lactate	0.643	<b>0.231</b>	0.324	<b>0.092</b>
GAM: (glucose: lactate)	<b>0.644</b>	<b>0.232</b>	<b>0.325</b>	<b>0.092</b>
APACHE-IVa	0.829	0.426	0.262	0.077

**Table 2.** Weighted cross-validation results (mean values) for all patients and stratified by diabetic status. Best metric values are in [bold].



**Figure 3.** GAM model partial effects (log odds scale) for several GAM models. **(A)** The impact of adjustment for lactate on the partial effect of glucose. **(B)** The partial effects of glucose for diabetics and non-diabetics GAM: *glucose + lactate*. **(C)** The 2D spline interaction effect between blood glucose and lactate levels for non-diabetics GAM: (*glucose:lactate* |  $DM=0$ ). **(D)** The 2D spline interaction effect between blood glucose and lactate levels for diabetics GAM: (*glucose:lactate* |  $DM=1$ ).

## Discussion

The aim of this research project was to investigate the relationship between blood glucose and hospital mortality while accounting for blood lactate measurements. Across 43,250 ICU admissions, weighted to account for missing data, we assessed the predictive ability of several models stratified by diabetic status. Additionally, we varied the functional form of the model, using binning or flexible semi-parametric GAMs to model the continuous variables. We found that inclusion of blood glucose improved predictive performance. Assessment of the functional form of the GAMs revealed that in non-diabetics hyperglycaemia remained a risk factor for hospital mortality with a lessened effect in diabetics. In both subgroups hypoglycaemia remained a risk. Sensitivity analyses using alternative models and approaches to accounting for missing data supported these findings.

**Relationship to previous literature.** The relationship between hyperglycaemia, hyperlactatemia and outcome in critically ill patients has been assessed in previous research. Two early studies found conflicting results on the impact of adjusting for lactate on the association between glucose and hospital outcome. Martin et al.<sup>7</sup> found both glucose and lactate on admission to be independent predictors of mortality in a study of 1,551 surgical ICU patients, with an 18 mg/dL increase in blood glucose associated with a 1% increase in the odds of mortality when controlling for lactate. In contrast, Kaukonen et al.<sup>6</sup> found that inclusion of lactate nullified the predictive power of glucose on hospital outcome in 7,925 mixed ICU patients. A common feature of these studies is the use of glucose as linear term in logistic regression. Later studies using discretisation—whether through stratification or binning—found an interactive effect between the two variables. Freire Jorge et al.<sup>8</sup> found similar results to the present research, with abnormal levels of both blood glucose and blood lactate associated with the poorest outcomes, a finding we reproduced (see Fig. 3C and LR coefficient table in Appendix A Table A6). Additionally in the current study, using two-dimensional spline terms, we found an interaction effect such that with increased blood lactate the lowest risk blood glucose level increased, from approximately 100 mg/dL to



150 mg/dL as blood lactate rose from 1 to 10 mmol/L (Fig. 3C). In an alternative approach Chen et al.<sup>5</sup> assessed the additional impact of lactate on glucose's prognostic value finding it added value for patients with hypo- and hyperglycaemia. Examining the results of the current study, these findings, while superficially contradictory, may be explained by variation in study design and analysis techniques as discussed further below.

The current results are in line with previous research suggesting that admission blood glucose has a different risk profile for diabetics and non-diabetics. It has been suggested that chronic exposure to hyperglycaemia in diabetes results in metabolic adaptation, reducing the potential toxic effects of acute hyperglycaemia in critical illness<sup>39</sup>. While no mechanism for this elevation has been described, higher HbA1c has been associated with an elevated renal threshold for glucose<sup>40</sup>, illustrating that diabetes may result in altered physiological homeostasis, albeit with significant negative side-effects<sup>41</sup>. Indeed, several authors have found a blunted or absent relationship between acute hyperglycaemia and mortality risk in diabetics<sup>42–45</sup>, in line with the current findings. On the other hand, there is evidence that the threshold for hypoglycaemic risk is elevated in those with uncontrolled diabetes<sup>46</sup>. In the current study we found that patients with diabetes have an inflection towards greater risk around 100–150 mg/dL (Fig. 3B), higher than standard definitions of hypoglycaemia, but in line with American Diabetes Association guidelines that a range of > 140 mg/dL is appropriate for hospitalised diabetics<sup>47</sup>.

**Implications of study findings.** Two features of the previous research on glucose and lactate stand out in relationship to the current study. None of the previous studies reported results either stratified by diabetic status or equivalently adjusted using interaction terms. As discussed above and shown in the current study (Fig. 3B), the relationship between blood glucose and mortality is modified by diabetic status. Given the “diabetic pandemic”, adjusting for diabetic status (or HbA1c level) when investigating the relationship between blood glucose and an outcome is clearly important.

In situations where risk profiles (or other physiological effects) may be non-linear, or interactive, model choice is important, and differences may explain variation in the previous results. As suggested by Freire Jorge et al.<sup>8</sup> the use of glucose as linear term may partially explain the results in Kaukonen et al.<sup>6</sup>. More generally, we found that the binned models underperformed the spline-based GAM models despite no loss in model interpretability (Table 2). At the extreme, choice of bin cut point to account for non-linear effects can lead to spurious results<sup>48</sup>. While purely non-linear methods such as deep learning are increasingly used to account for non-linear effects, such models remain largely uninterpretable, with interpretability metrics not even guaranteed to give similar findings<sup>49</sup>. Thus these models, as currently implemented, are not ideal for informing high stakes decisions such as those made in medicine<sup>50</sup>, or for answering certain scientific questions. GAMs and other “white-box” machine learning approaches<sup>51</sup> present an alternative approach for situations where non-linear effects are expected and interpretability is key.

**Strengths and weaknesses.** This paper has several strength and weakness. Strengths noted above are comparison of several model choices. The predictive aspect of model selection enables easy comparison of models with complex interactive effects and different estimation methods<sup>52</sup>. Through varying the model form and assessing predictive ability overall and across subgroups we gain a greater idea of the true discriminatory power of the models in the studied population. Other strengths include the use of a large high quality real world data collection<sup>24,25</sup>. The data source consists of a mixed ICU population across a wide geographic area, with potential variation in treatment practices and cohorts treated. However, the retrospective observational nature of the data source is also a weakness. The data were not originally designed for use in this study, and a significant portion of patients in the source database did not have a blood lactate measurement or the glucose and lactate measurements were separated by several hours. Additionally, we did not attempt to disentangle the degree to which exogenous glucose was responsible for measured blood glucose levels. Missing data is a known cause of bias in analysis of EMR data, with sicker patients typically have more complete records, as seen in the current findings<sup>53,54</sup>. While inverse probability of missingness weighting offers a straightforward approach to designing analyses aimed at reducing this bias the weighting did not completely remove differences between the groups (Table 1), and fundamentally cannot account for unmeasured predictors of missingness.

**Future directions.** While this research has confirmed that blood glucose is a marker of poor prognosis even in the presence of stronger markers (blood lactate) further research is required into the potential mechanisms of any causal effect on patient outcome along with subgroup variation along with examination of alternative outcomes, such as organ failure<sup>55</sup>. Ultimately lactate is produced from glucose, and discounting the presence of tissue hypoxia<sup>14</sup>, both hyperglycaemia and hyperlactatemia are measures of altered energy metabolism. However, biophysiological theories of how acute stress disturbances in energy metabolism cause poor outcomes are only starting to be pieced together<sup>9,16</sup>. Given the long timeframes of RCTs<sup>56</sup> future research should continue aiming to link theory and observational data. For instance, if as claimed by Gunst et al.<sup>17</sup> the treatment effect of insulin is entirely through reducing glucose toxicity, it should be possible to demonstrate this in observational data using methods to estimate causal effects of time-varying exposures that treat blood glucose measures as mediators such as g-computation<sup>57</sup>. Previous research looking at the treatment effect of insulin in observational data has found a negative impact of insulin treatment<sup>58</sup>, largely explained by huge differences in rates of hypoglycaemia (29% vs 1.4%) between treated and non-treated. Causal inference in observational data is notoriously difficult and a wider range of study designs need to be applied to the data to investigate this issue, in particular looking at subgroup effects (e.g. diabetic vs non-diabetic or surgical vs non-surgical patients<sup>17</sup>).

## Conclusion

In a mixed ICU population admission blood glucose is predictive of hospital mortality after accounting for blood lactate, with different hyperglycaemic risk profiles for diabetics and non-diabetics. In diabetics we found no association between hyperglycaemia and hospital mortality, while in non-diabetics hyperglycaemia remains a predictor of poor outcomes.

## Data availability

The data underlying this article are freely available at <https://eicu-crd.mit.edu/> and can be accessed following completion of the required training and data usage agreements.

Received: 19 May 2023; Accepted: 18 September 2023

Published online: 21 September 2023

## References

- Mongkolpun, W., Provenzano, B. & Preiser, J.-C. Updates in glycemic management in the hospital. *Curr. Diabetes Rep.* **19**(11), 1–6 (2019).
- Kruse, O., Grunnet, N. & Barfod, C. Blood lactate as a predictor for in-hospital mortality in patients admitted acutely to hospital: A systematic review. *Scand. J. Trauma Resusc. Emerg. Med.* **19**(1), 1–12 (2011).
- Cichosz, S. L. & Schaarup, C. Hyperglycemia as a predictor for adverse outcome in ICU patients with and without diabetes. *J. Diabetes Sci. Technol.* **11**(6), 1272–1273 (2017).
- Cichosz, S. L., Redke, F. & Hejlesen, O. K. Spontaneous and iatrogenic hypoglycaemia related to mortality in the ICU. *Diabetes Metab.* **45**(6), 545–549 (2019).
- Chen, X., Bi, J. & Wu, R. The impact of serum glucose on the predictive value of serum lactate for hospital mortality in critically ill surgical patients. *Dis. Mark.* **2019** (2019).
- Kaukonen, K.-M. *et al.* Stress hyperlactatemia modifies the relationship between stress hyperglycemia and outcome: A retrospective observational study. *Crit. Care Med.* **42**(6), 1379–1385 (2014).
- Martin, J. *et al.* Point-of-care testing on admission to the intensive care unit: Lactate and glucose independently predict mortality. *Clin. Chem. Lab. Med. (CCLM)* **51**(2), 405–412 (2013).
- Freire Jorge, P. *et al.* The association of early combined lactate and glucose levels with subsequent renal and liver dysfunction and hospital mortality in critically ill patients. *Crit. Care* **21**(1), 1–11 (2017).
- Brooks, G. A. The science and translation of lactate shuttle theory. *Cell Metab.* **27**(4), 757–785 (2018).
- Fitzgerald, O. *et al.* Incorporating real-world evidence into the development of patient blood glucose prediction algorithms for the ICU. *J. Am. Med. Inform. Assoc.* **28**, 1642–1650 (2021).
- Vedantam, D., Poman, D. S., Motwani, L., Asif, N., Patel, A. & Anne, K. K. Stress-induced hyperglycemia: Consequences and management. *Cureus* **14**(7) (2022).
- Becker, C. D. *et al.* Hyperglycemia in medically critically ill patients: Risk factors and clinical outcomes. *Am. J. Med.* **133**(10), e568–e574 (2020).
- Ferreruela, M., Raurich, J. M., Ayestarán, I. & Llopart-Pou, J. A. Hyperlactatemia in ICU patients: Incidence, causes and associated mortality. *J. Crit. Care* **42**, 200–205 (2017).
- García-Alvarez, M., Marik, P. & Bellomo, R. Stress hyperlactatemia: Present understanding and controversy. *Lancet Diabetes Endocrinol.* **2**(4), 339–347 (2014).
- James, J. H., Luchette, F. A., McCarter, F. D. & Fischer, J. E. Lactate is an unreliable indicator of tissue hypoxia in injury or sepsis. *Lancet* **354**(9177), 505–508 (1999).
- Nolt, B. *et al.* Lactate and immunosuppression in sepsis. *Shock (Augusta, Ga.)* **49**(2), 120 (2018).
- Gunst, J., De Bruyn, A. & Van den Berghe, G. Glucose control in the ICU. *Curr. Opin. Anaesthesiol.* **32**(2), 156 (2019).
- Marik, P. E. Precision glycemic control in the ICU. *Crit. Care Med.* **44**(7), 1433–1434 (2016).
- Griesdale, D. E., de Souza, R. J. & Finfer, S. Intensive insulin therapy and mortality among critically ill patients: A meta-analysis including NICE-SUGAR study data. *Cmaj* **180**(8), 821–827 (2009).
- Van den Berghe, G. How does blood glucose control with insulin save lives in intensive care?. *J. Clin. Investig.* **114**(9), 1187–1195 (2004).
- Van Niekerk, G., Davis, T. & Engelbrecht, A.-M. Hyperglycaemia in critically ill patients: The immune system's sweet tooth. *Crit. Care* **21**, 1–5 (2017).
- Ivashkiv, L. B. The hypoxia–lactate axis tempers inflammation. *Nat. Rev. Immunol.* **20**(2), 85–86 (2020).
- Wood, S. N. Fast stable direct fitting and smoothness selection for generalized additive models. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **70**(3), 495–518 (2008).
- Pollard, T. J. *et al.* The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci. Data* **5**(1), 1–13 (2018).
- O'Halloran, H. M., Kwong, K., Veldhoen, R. A. & Maslove, D. M. Characterizing the patients, hospitals, and data quality of the eICU collaborative research database. *Crit. Care Med.* **48**(12), 1737–1743 (2020).
- Fitzgerald, O., Perez-Concha, O., Gallego Luxan, B., Rudd, L. & Jorm, L. Curation and description of a blood glucose management and nutritional support cohort using the eICU collaborative research database. *medRxiv* 2023.04.20.23288845 (2023).
- Wickham, H. & Bryan, J. bigquery: An Interface to Google's BigQuery™ API. R package version. **1**(0) (2018).
- Dowle, M. & A. Srinivasan, *Data. Table: Extension of Data. Frame. R Package Version 1.12. 2*, ed (2019).
- Von Elm, E. *et al.* The strengthening of reporting of observational studies in epidemiology (STROBE) statement: Guidelines for reporting observational studies. *Lancet* **370**(9596), 1453–1457 (2007).
- Li, L., Shen, C., Li, X. & Robins, J. M. On weighting approaches for missing data. *Stat. Methods Med. Res.* **22**(1), 14–30 (2013).
- Zimmerman, J. E., Kramer, A. A., McNair, D. S. & Malila, F. M. Acute Physiology and Chronic Health Evaluation (APACHE) IV: Hospital mortality assessment for today's critically ill patients. *Crit. Care Med.* **34**(5), 1297–1310 (2006).
- Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2016).
- Chernozhukov, V. *et al.* *Double/Debiased Machine Learning for Treatment and Structural Parameters* (Oxford University Press, 2018).
- Wood, S. N. Thin plate regression splines. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **65**(1), 95–114 (2003).
- Markatou, M., Afendras, G. & Agostinelli, C. Weighted cross validation in model selection. *Wiley Interdiscip. Rev. Comput. Stat.* **10**(6), e1439 (2018).
- RR Core Team. *R: A Language and Environment for Statistical Computing* (2013).
- Wickham, H. ggplot2. *Wiley Interdiscip. Rev. Comput. Stat.* **3**(2), 180–185 (2011).

38. Simpson, G. L. & Singmann, H. *Package gratia*. Ggplot-based graphics and other useful functions for GAMs fitted using MgcV, 0.1-0 (Ggplot-based graphics and utility functions for working with GAMs fitted using the mgcv package). [Google Scholar] (2018).
39. Siegelaar, S. E., Hoekstra, J. B. & DeVries, J. H. Special considerations for the diabetic patient in the ICU; targets for treatment and risks of hypoglycaemia. *Best Pract. Res. Clin. Endocrinol. Metab.* **25**(5), 825–834 (2011).
40. Hieshima, K., Sugiyama, S. & Jinnouchi, T. Elevation of the renal threshold for glucose is associated with insulin resistance and higher glycated hemoglobin levels. *J. Diabetes Investig.* **11**(3), 617–625 (2020).
41. Giri, B. *et al.* Chronic hyperglycemia mediated physiological alteration and metabolic distortion leads to organ dysfunction, infection, cancer progression and other pathophysiological consequences: An update on glucose toxicity. *Biomed. Pharmacother.* **107**, 306–328 (2018).
42. Egi, M. *et al.* Blood glucose concentration and outcome of critical illness: The impact of diabetes. *Crit. Care Med.* **36**(8), 2249–2255 (2008).
43. Sechterberger, M. K. *et al.* The effect of diabetes mellitus on the association between measures of glycaemic control and ICU mortality: A retrospective cohort study. *Crit. Care* **17**(2), 1–10 (2013).
44. van Vught, L. A., Holman, R., de Jonge, E., de Keizer, N. F. & Van der Poll, T. Diabetes is not associated with increased 90-day mortality risk in critically ill patients with sepsis. *Crit. Care Med.* **45**(10), e1026–e1035 (2017).
45. Krinsley, J. S., Egi, M. & Mackenzie, I. M. Diabetic status and the relation of the three domains of glycemic control to mortality in critically ill patients: an international multicenter cohort study. *Crit. Care* **17**(2), 1–17 (2013).
46. Egi, M., Krinsley, J. S. & Bellomo, R. Pre-morbid glycemic control modifies the interaction between acute hypoglycemia and mortality. *Intensive Care Med.* **42**(4), 562–571 (2016).
47. American Diabetes Association. 15. Diabetes care in the hospital: Standards of medical care in diabetes-2019. *Diabetes Care* **42**(Suppl 1), S173–S181 (2019).
48. Royston, P., Altman, D. G. & Sauerbrei, W. Dichotomizing continuous predictors in multiple regression: A bad idea. *Stat. Med.* **25**(1), 127–141 (2006).
49. Krishna, S., Han, T., Gu, A., Pombra, J., Jabbari, S., Wu, S. & Lakkaraju, H. *The Disagreement Problem in Explainable Machine Learning: A Practitioner's Perspective*. arXiv preprint <http://arxiv.org/abs/2202.01602> (2022).
50. Rudin, C. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nat. Mach. Intell.* **1**(5), 206–215 (2019).
51. Loyola-Gonzalez, O. Black-box vs. white-box: Understanding their advantages and weaknesses from a practical point of view. *IEEE Access* **7**, 154096–154113 (2019).
52. Geisser, S. & Eddy, W. F. A predictive approach to model selection. *J. Am. Stat. Assoc.* **74**(365), 153–160 (1979).
53. Rusanov, A., Weiskopf, N. G., Wang, S. & Weng, C. Hidden in plain sight: Bias towards sick patients when sampling patients with sufficient electronic health record data for research. *BMC Med. Inform. Decis. Mak.* **14**(1), 1–9 (2014).
54. Weber, G. M., Adams, W. G. & Murphy, S. N. Biases introduced by filtering electronic health records for patients with “complete data”. *J. Am. Med. Inform. Assoc.* **24**(6), 1134–1141 (2017).
55. Pretty, C. G. *et al.* Variability of insulin sensitivity during the first 4 days of critical illness: Implications for tight glycemic control. *Ann. Intensive Care* **2**(1), 1–10 (2012).
56. Black, N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ* **312**(7040), 1215–1218 (1996).
57. Robins, J. A new approach to causal inference in mortality studies with a sustained exposure period—Application to control of the healthy worker survivor effect. *Math. Model.* **7**(9–12), 1393–1512 (1986).
58. Yu, B., Li, C., Sun, Y. & Wang, D. W. Insulin treatment is associated with increased mortality in patients with COVID-19 and type 2 diabetes. *Cell Metabol.* **33**(1), 65–77 (2021).

## Author contributions

O.F. conceived the study and carried out the analysis, and drafted and refined the manuscript. All authors provided critical feedback throughout the study and helped shape the manuscript.

## Funding

This work was supported by the Commonwealth Industrial and Scientific Research Organisation (CSIRO), eHealth NSW, and the Australian government through an Australian Government Research Training Program scholarship and CSIRO top up scholarship.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-43044-7>.

**Correspondence** and requests for materials should be addressed to O.F.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023