Supplementary Appendix 1. Candidate Predictors for Inpatient Hypoglycemia Prediction

Table 1. Candidate predictors included in model development, with definitions, units, missing data, and imputation strategies.

Category	Variable	Data in EHR	Data in Models	Unit	Missing (%)	Imputation Strategy
Demographics	Sex	Patient sex	Binary (1=Male)	N/A	0.0%	N/A
	Smoking	History of smoking	Binary (1=Yes)	N/A	0.5%	Mode / Baseline
	Drinking	History of alcohol use	Binary (1=Yes)	N/A	0.5%	Mode / Baseline
	Age	Age at admission	Continuous	years	0.0%	N/A
	Weight	Body weight at admission	Continuous	kg	0.5%	Median/Baseline
	Height	Height at admission	Continuous	cm	0.5%	Median/Baseline
	BMI	Body mass index	Continuous	kg/m^2	0.6%	Median/Baseline
	T1DM	Type 1 Diabetes	Binary (1=T1DM)	N/A	0.0%	N/A
Comorbidities	HTN	Hypertension (ICD-10 I10)	Binary (1=Yes)	N/A	0.0%	N/A
	HLD	Hyperlipidemia (ICD-10 E78)	Binary (1=Yes)	N/A	0.0%	N/A
	CAD	Coronary artery disease	Binary (1=Yes)	N/A	0.0%	N/A
	Malignancy	Malignant tumors	Binary (1=Yes)	N/A	0.0%	N/A
	CRF	Chronic renal failure	Binary (1=Yes)	N/A	0.0%	N/A
	RRT	Renal replacement therapy	Binary (1=Yes)	N/A	0.0%	N/A
	DPVD	Diabetic peripheral vascular disease	Binary (1=Yes)	N/A	0.0%	N/A
	DPN	Diabetic peripheral neuropathy	Binary (1=Yes)	N/A	0.0%	N/A
	DF	Diabetic foot	Binary (1=Yes)	N/A	0.0%	N/A
	DN	Diabetic nephropathy	Binary (1=Yes)	N/A	0.0%	N/A
	DR	Diabetic retinopathy	Binary (1=Yes)	N/A	0.0%	N/A
Laboratory	TBIL	Total bilirubin	Continuous	μmol/L	19.14%	Multiple Imputation
	TC	Total cholesterol	Continuous	mmol/L	78.4%	Excluded
	TG	Triglycerides	Continuous	mmol/L	76.8%	Excluded
	LDL	Low-density lipoprotein	Continuous	mmol/L	76.8%	Excluded
	HDL	High-density lipoprotein	Continuous	mmol/L	75.9%	Excluded
	Na^+	Serum sodium	Continuous	mmol/L	18.01%	Multiple Imputation
	K^+	Serum potassium	Continuous	mmol/L	18.04%	Multiple Imputation
	RBC	Red blood cell count	Continuous	$\times 10^{12}/L$	17.89%	Multiple Imputation
	WBC	White blood cell count	Continuous	$\times 10^9/L$	17.89%	Multiple Imputation
	PLT	Platelet count	Continuous	$\times 10^9/L$	17.93%	Multiple Imputation
	Hb	Hemoglobin	Continuous	g/L	17.89%	Multiple Imputation
	hs-CRP	High-sensitivity C-reactive protein	Continuous	mg/L	76.07%	Excluded
	HbA1c	Glycated hemoglobin	Continuous	%	19.7%	Multiple Imputation
	ALT	Alanine aminotransferase	Continuous	U/L	19.23%	Multiple Imputation
	AST	Aspartate aminotransferase	Continuous	U/L	19.71%	Multiple Imputation
	GGT	Gamma-glutamyl transferase	Continuous	U/L	19.95%	Multiple Imputation
	ALB	Albumin	Continuous	g/L	19.06%	Multiple Imputation
	SCr	Serum creatinine	Continuous	μmol/L	18.46%	Multiple Imputation
	eGFR	Estimated glomerular filtration rate	Continuous	mL/min/1.73 m^2	18.95%	Multiple Imputation
Antidiabetic Medications	Metformin	Medication orders	Binary (1=Yes)	N/A	0.0%	N/A
	SUs	Sulfonylureas	Binary (1=Yes)	N/A	0.0%	N/A
	Glinides	Glinides	Binary (1=Yes)	N/A	0.0%	N/A
	TZDs	Thiazolidinediones	Binary (1=Yes)	N/A	0.0%	N/A
	AGIs	Alpha-glucosidase inhibitors	Binary (1=Yes)	N/A	0.0%	N/A
	DPP-4i	DPP-4 inhibitors	Binary (1=Yes)	N/A	0.0%	N/A
	SGLT2i	SGLT2 inhibitors	Binary (1=Yes)	N/A	0.0%	N/A
	Bolus Insulin	Mealtime insulin	Binary (1=Yes)	N/A	0.0%	N/A
	Premixed Insulin	Premixed insulin	Binary (1=Yes)	N/A	0.0%	N/A
	Dual Insulin	Basal + bolus insulin	Binary (1=Yes)	N/A	0.0%	N/A
	Basal Insulin	Basal insulin only	Binary (1=Yes)	N/A	0.0%	N/A
Vital Signs	T	First recorded temperature on prior day	Continuous	°C	0.7%	Multiple Imputation
	P	First recorded pulse	Continuous	beats/min	0.8%	Multiple Imputation
	RR	First recorded respiratory rate	Continuous	breaths/min	0.7%	Multiple Imputation

	SBP	Systolic blood pressure	Continuous	mmHg	0.3%	Multiple Imputation
	DBP	Diastolic blood pressure	Continuous	mmHg	0.3%	Multiple Imputation
Glycemic Metrics	INPMEAN	Mean glucose since admission	Continuous	mmol/L	3.2%	Forward fill /Multiple Imputation
	INPCV	Glucose coefficient of variation since admission	Continuous	%	6.8%	Forward fill /Multiple Imputation
	24HMIN	Minimum glucose prior day	Continuous	mmol/L	3.2%	Forward fill /Multiple Imputation
	24HMAX	Maximum glucose prior day	Continuous	mmol/L	3.2%	Forward fill /Multiple Imputation
	24HMEAN	Mean glucose prior day	Continuous	mmol/L	3.2%	Forward fill /Multiple Imputation
	24HCV	Glucose coefficient of variation prior day	Continuous	%	3.2%	Forward fill /Multiple Imputation
Hospitalization	PREHYPO	Any previous hypoglycemia	Binary (1=Yes)	N/A	0.0%	N/A
	HospDay	Calendar day index of hospitalization	Continuous	days	0.1%	Forward fill
	NPO	Fasting status on prior day	Binary (1=Yes)	N/A	0.0%	N/A

Abbreviations: BMI = body mass index; mmol/L = millimoles per liter; μmol/L = micromoles per liter; g/L = grams per liter; U/L = units per liter; CV = coefficient of variation.

Note: “Data in EHR” indicates the original source or measurement method in electronic health records. “Data in Models” indicates how the variable was represented or aggregated for modeling. Missing (%) is the proportion of missing values in the dataset. Imputation Strategy: Multiple Imputation = missing values imputed using multiple imputation methods; Mode / Baseline = imputed with baseline data or the most frequent category; Excluded (>25%) = variables with >25% missing data were excluded from modeling.

Table 2. Time-series glucose dynamics features used in model development, with formulas, time windows, units, and notes.9

Feature Category	Formula	Time Window	Notes
Change	$\text{glucose\_change} = \frac{\boldsymbol{G_t - G_{t-1}}}{\boldsymbol{G_{t-1}}}$	24h / 7d	Change of mean glucose (daily / weekly / monthly)
Amplitude	$\text{glucose\_amplitude} = \mathbf{max}(\boldsymbol{G}) - \mathbf{min}(\boldsymbol{G})$	24h / 7d	Peak-to-trough amplitude
Peak-to-trough ratio	$\text{glucose\_ratio} = \frac{\mathbf{max}(\boldsymbol{G})}{\mathbf{min}(\boldsymbol{G})}$	24h / 7d	Daily / weekly / monthly peak-to-trough ratio
Linear trend	glucose_trend = slope of polyfit(range, G, 1)	24h / 7d	Linear trend over time window
Volatility	$\text{glucose\_volatility} = \frac{\sigma(\boldsymbol{G})}{\overline{\boldsymbol{G}}}$	24h / 7d	Coefficient of variation
Time-decay weighted mean	$\text{time\_weighted} = \frac{\sum_{i=1}^{N} \boldsymbol{G_i} \cdot \boldsymbol{e^{-i/N}}}{\sum_{i=1}^{N} \boldsymbol{e^{-i/N}}}$	24h / 7d	Exponentially decayed average
Quadratic trend	time_trend = quadratic coefficient of polyfit(range, G, 2)	24h / 7d	Captures curvature over time
Main frequency component	$\text{time\_periodic} = \text{FFT}(\boldsymbol{G})_{\mathbf{1}}$	24h / 7d	Dominant frequency in glucose signal
Time-weighted average	$\text{glucose\_twa} = \frac{\sum_i \boldsymbol{G_i} \cdot \boldsymbol{w_i}}{\sum_i \boldsymbol{w_i}}, \quad \boldsymbol{w_i} = \boldsymbol{e^{-i/N}}$	24h / 7d	Weighted mean with time decay
Stability index	$\text{stability\_index} = \frac{\mathbf{1}}{\mathbf{1} + \sigma(\boldsymbol{G})}$	24h / 7d	Higher = more stable
Volatility index	$\text{volatility\_index} = \frac{\sigma(\boldsymbol{G})}{\overline{\boldsymbol{G}}}$	24h / 7d	Normalized variability
Log-transformed mean	$\text{log\_glucose} = \mathbf{log}(\overline{\boldsymbol{G}} + \mathbf{1})$	24h	Daily log-transformed mean

Abbreviations: FFT = fast Fourier transform; TWA = time-weighted average.

Note: All temporal glucose metrics were calculated based on the preceding 24 hours, or 7 days windows, as indicated. $\boldsymbol{G}$= glucose values in the time window $\overline{\boldsymbol{G}}$ = mean glucose $\sigma(\boldsymbol{G})$= standard deviation of glucose; $\boldsymbol{N}$ = number of time points in the window.