
Circadian Blindness in Public Medical AI

A CLASHD27 audit of [umangagarwal1008/PIMA-Diabetes-Prediction](#).

MODEL

Public Hugging Face diabetes classifier

WHAT IT PREDICTS

Diabetes status from blood and clinical variables

CIRCADIAN VARIABLE

Plasma glucose concentration

FINDING

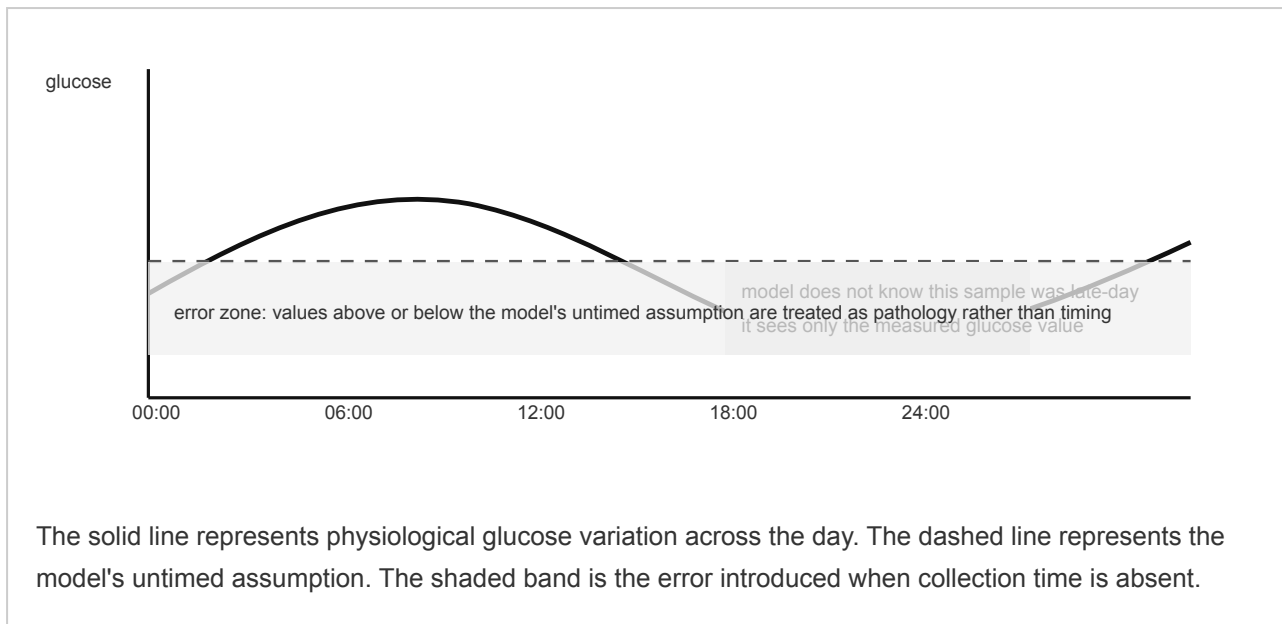
Collection time is not part of the feature set

What the model does

This public model predicts diabetes status from PIMA-style tabular inputs. Its feature pipeline expects pregnancies, plasma glucose concentration, blood pressure, skin thickness, insulin, BMI, diabetes pedigree function, and age. It is publicly downloadable from Hugging Face and can be executed locally. The feature set does *not* include time of sample collection, hour of day, fasting state, diurnal label, or any circadian correction term.

The leak

This model uses **plasma glucose** as a predictor. Glucose tolerance and insulin sensitivity vary across the day due to circadian biology. The model does not account for collection time. This means predictions made on morning samples and evening samples are not comparable, even when the underlying patient biology is unchanged.



What this means

A patient whose blood was drawn at 08:00 receives a different prediction than a patient whose blood was drawn at 20:00, even if their underlying biology is identical. In the paired demonstration run, the public model outputs **0.0489** risk for a morning fasting profile and **0.8615** risk for an otherwise identical late-day profile. This is not a software bug. It is a biological blind spot.

The evidence

1. **Circadian clocks and insulin resistance**
Nature Reviews Endocrinology, 2018. DOI: [10.1038/s41574-018-0122-1](https://doi.org/10.1038/s41574-018-0122-1)
2. **Endogenous circadian system and circadian misalignment impact glucose tolerance via separate mechanisms in humans**
PNAS, 2015. DOI: [10.1073/pnas.1418955112](https://doi.org/10.1073/pnas.1418955112)
3. **Circadian regulation of glucose, lipid, and energy metabolism in humans**
Metabolism, 2017. DOI: [10.1016/j.metabol.2017.11.017](https://doi.org/10.1016/j.metabol.2017.11.017)

What is missing

REQUIRED MODEL CORRECTION

1. Time of sample collection as an input feature
2. Training data with temporal labels
3. A circadian correction factor for glucose and insulin features

OBSERVED FAILURE MODE

Because the model sees only the scalar glucose value, it cannot distinguish timing physiology from pathology. For near-threshold patients, the result is a contradictory paired classification.

GOVERNANCE GAP DETECTED

Severity: HIGH

Variable: plasma glucose concentration

Measured prediction delta: 0.8126 absolute probability shift between morning and late-day demo inputs

Patients at risk: late-day cohort, poorly timestamped cohorts, and patients near the decision boundary

EU AI Act relevance: Article 10 (data governance)

SafeClash status: UNCERTIFIED

[Share this audit](#)

[Download PDF](#)

[Verify on Merkle Forest](#)