

<https://doi.org/10.1038/s41746-025-01709-9>

A distributional reinforcement learning model for optimal glucose control after cardiac surgery



Jacob M. Desman^{1,2}, Zhang-Wei Hong³, Moein Sabounchi^{1,2}, Ashwin S. Sawant^{1,2,4}, Jaskirat Gill⁵, Ana C. Costa⁶, Gagan Kumar⁷, Rajeev Sharma⁸, Arpeta Gupta⁹, Paul McCarthy¹⁰, Veena Nandwani¹⁰, Doug Powell¹⁰, Alexandra Carideo⁵, Donnie Goodwin¹⁰, Sanam Ahmed⁵, Umesh Gidwani⁵, Matthew A. Levin¹¹, Robin Varghese^{5,6}, Farzan Filsoufi⁶, Robert Freeman¹, Avniel Shetreat-Klein¹², Alexander W. Charney¹, Ira Hofer^{1,2,11}, Lili Chan^{1,2,13}, David Reich¹¹, Patricia Kovatch¹⁴, Roopa Kohli-Seth⁵, Monica Kraft¹⁵, Pulkit Agrawal³, John A. Kellum¹⁶, Girish N. Nadkarni^{1,2,13,17} & Ankit Sakhuja^{1,2,5,17} ✉

This study introduces Glucose Level Understanding and Control Optimized for Safety and Efficacy (GLUCOSE), a distributional offline reinforcement learning algorithm for optimizing insulin dosing after cardiac surgery. Trained on 5228 patients, tested on 920, and externally validated on 649, GLUCOSE achieved a mean estimated reward of 0.0 [−0.07, 0.06] in internal testing and −0.63 [−0.74, −0.52] in external validation, outperforming clinician returns of −1.29 [−1.37, −1.20] and −1.02 [−1.16, −0.89]. In multi-phase human validation, GLUCOSE first showed a significantly lower mean absolute error (MAE) in insulin dosing, with 0.9 units MAE versus clinicians' 1.97 units ($p < 0.001$) in internal testing and 1.90 versus 2.24 units ($p = 0.003$) in external validation. The second and third phases found GLUCOSE's performance as comparable to or exceeding that of senior clinicians in MAE, safety, effectiveness, and acceptability. These findings suggest GLUCOSE as a robust tool for improving postoperative glucose management.

Cardiac surgery elicits a substantial metabolic stress response resulting in postoperative hyperglycemia regardless of diabetic status¹. Post-operative hyperglycemia after cardiac surgery is common, occurring in 60–80% patients with diabetes², and over 50% non-diabetic patients³. It is associated with higher rates of post-operative infections^{4,5}, acute kidney injury^{3,6–8}, cardiac arrhythmias³, longer length of stay³, and higher mortality^{6–9}. Due to its significance, the Society of Thoracic Surgeons (STS) recommends maintaining blood glucose levels below 180 mg/dL after cardiac surgery¹⁰.

Achieving adequate glucose control post-operatively is challenging. A study found that only 15% of patients had appropriate glucose control, defined as glucose level between 70 mg/dL to 180 mg/dL, within the first day after cardiac surgery¹¹. This early post-operative period, when patients are critically ill and require care in intensive care unit (ICU) settings, is highly dynamic with rapidly changing clinical characteristics of patients. Currently, post-operative glucose management involves titration of regular insulin based on hospital specific protocols and the experience of treating clinicians.

¹The Charles Bronfman Institute for Personalized Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ²Division of Data-Driven and Digital Medicine, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ³Improbable AI Lab, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁴Division of Hospital Medicine, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁵Institute for Critical Care Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁶Department of Cardiothoracic Surgery, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ⁷Department of Pulmonary and Critical Care Medicine, Northeast Georgia Medical Center, Gainesville, GA, USA. ⁸Division of Endocrinology, Hackensack University Medical Center, Hackensack, NJ, USA. ⁹Division of Endocrinology, Millenium Physician Group, Jacksonville, FL, USA. ¹⁰Section of Cardiovascular Critical Care, Department of Cardiovascular and Thoracic Surgery, West Virginia University, Morgantown, WV, USA. ¹¹Department of Anesthesiology, Perioperative, and Pain Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ¹²Department of Rehabilitation and Physical Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ¹³Division of Nephrology, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ¹⁴Scientific Computing, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ¹⁵Samuel Bronfman Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA. ¹⁶Department of Critical Care Medicine, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. ¹⁷These authors contributed equally: Girish N. Nadkarni, Ankit Sakhuja. ✉e-mail: ankit.sakhuja@mssm.edu

However, due to the highly dynamic nature of this early post-operative period, some treatment regimens may be more suitable for certain patients or only effective for a limited time as their condition evolves. This leads to high rates of hyperglycemic and hypoglycemic episodes^{11,12}, both associated with worse outcomes, as these protocolized regimens often fail to account for individual patient variability in real-world settings^{13,14}. Therefore, personalized and dynamic insulin titration is crucial for improving glucose control in patients following cardiac surgery.

Prior algorithmic approaches to inpatient insulin management have primarily involved institution specific sliding scale doses, focused on glucose prediction, or used static daily insulin dose estimation. Sliding scale insulin regimens are standard across most institutions, but they are reactive and non-personalized, providing the same dose for a given glucose regardless of patient-specific factors, a practice that can be both ineffective and dangerous¹⁵. Nguyen et al. developed a supervised machine learning model to predict the total daily dose of insulin to improve upon weight-based dosing guidelines¹⁶. However, this approach excluded ICU patients and does not provide real-time dosing recommendations. Alternatively, while there exist many supervised machine learning models for inpatient glucose prediction to address such challenges in glycemic management, including in ICU settings, these models forecast glucose trends rather than recommending sequential insulin dosing strategies^{17,18}.

Several modeling approaches have been explored for glucose prediction and control, including stochastic modeling frameworks that leverage variable-length time-stamped data to capture seasonal glucose patterns. For example, a seasonal stochastic local modeling approach (Glucose Prediction under Variable-Length Time-Stamped Daily Events) has been proposed to address inter-day variability in glucose regulation¹⁹. While these models offer valuable predictive capabilities, they often lack adaptive decision-making mechanisms for real-time insulin dosing. In contrast, reinforcement learning (RL) provides a dynamic approach that learns optimal insulin dosing policies by maximizing cumulative rewards in response to patient-specific glucose fluctuations. By transitioning from predictive modeling to decision-based RL frameworks, we aim to enhance personalized glucose management in the high-risk postoperative setting.

RL is a type of machine learning where an agent learns to make decisions by performing actions in an environment to maximize cumulative rewards²⁰. RL algorithms receive feedback in the form of rewards or penalties based on the actions taken, allowing the agent to improve its policy over time. This adaptability makes RL particularly well-suited for tasks that involve complex decision-making and require real-time adjustments, such as insulin titration in the dynamic postoperative environment.

Implementing an RL-based system for insulin titration can address the limitations of current glucose management protocols. By continuously learning from individual patient data, RL can provide personalized treatment plans that account for specific patient variability and maintain glucose in optimal range. Additionally, RL's capability to adapt to rapidly changing clinical characteristics ensures that insulin dosing remains optimal as patient conditions evolve. Traditionally, offline RL, where the agent learns from a fixed dataset without further interaction with the environment, has been limited by its focus on expected rewards, often overlooking the uncertainty in patient responses²¹. This limitation can lead to suboptimal treatment plans, as it fails to account for the full spectrum of possible outcomes. As a result, offline RL systems may not adequately address the diverse risk profiles associated with different patient actions, potentially compromising the safety and effectiveness of interventions²².

Our approach addresses this limitation by integrating distributional RL, which characterizes the entire distribution of potential outcomes rather than just the expected reward²². This methodology provides a more comprehensive understanding of the risks and benefits associated with various actions, allowing for more nuanced decision-making under uncertainty^{22–24}. By considering the full range of potential patient responses, distributional RL can enhance the personalization and safety of insulin titration protocols, ensuring optimal dosing as patient conditions change.

Our proposed model, Glucose Level Understanding and Control Optimized for Safety and Efficacy (GLUCOSE), aims to improve glucose management on the first day after cardiac surgery, potentially leading to better patient outcomes and more effective clinical decision-making. We have developed GLUCOSE using data of patients undergoing cardiac surgery in the Medical Information Mart for Intensive Care-IV (MIMIC-IV) database²⁵. We then validated the model externally with cardiac surgery patients from the eICU Collaborative Research Database (eICU-CRD), a diverse, multicenter database of critically ill patients²⁶.

Results

Study population

GLUCOSE was trained and validated on two separate ICU datasets: the MIMIC-IV database²⁵ and the eICU-CRD database²⁶. MIMIC-IV was used as the development cohort and split into training and internal testing sets. eICU-CRD was used as the external validation dataset. Based on the inclusion and exclusion criteria, our study included 6,148 patients in development dataset and 649 patients in external validation dataset. The mean age of patients in the development dataset was 67.8 ± 11.6 years with 71.1% males, and in external validation dataset was 67.0 ± 11.3 years with 67.2% males. At least one hypoglycemic event (< 70 mg/dL) occurred in 7.6% of patients in the development dataset and among 7.2% of patients in the external validation dataset. Similarly, at least one hyperglycemic event (> 180 mg/dL) occurred in 47.8% of patients in development dataset and 47.3% of patients in external validation dataset. The baseline characteristics of the patients are shown in Table 1 and Supplementary Table 1. The overall structure of our study is illustrated in Fig. 1.

Performance of GLUCOSE

To mitigate the sampling and stochastic biases inherent in offline RL²⁷ we trained, in line with previous literature, multiple models until we observed no significant improvements in the RL policies^{28–30}. Consequently, we trained 200 independent models and selected the model with the maximal lower bound of the 95% CI of mean estimated performance returns within the internal testing set as the GLUCOSE model. We compared the estimated performance returns of GLUCOSE at the lower bound of its 95% CI with the upper bound of clinicians' 95% CI (Fig. 2a) using fitted Q estimation (FQE) for off policy evaluation (OPE)^{31,32}, illustrating the differences in average estimated performance after evaluating 200 policies. The dotted blue and dotted orange lines reflect the 95% confidence intervals of the mean performance for the observed clinicians behavior in internal testing and external validation, respectively, while their non-dotted counterparts reflect the estimated performance of GLUCOSE through OPE (Fig. 2a). The best model, GLUCOSE, resulted in a mean estimated performance return of 0.0 [−0.07, 0.06] in the internal testing set and −0.63 [−0.74, −0.52] in the external validation dataset, showing significant improvements over the clinician returns of −1.29 [−1.37, −1.20] in the internal testing set and −1.02 [−1.16, −0.89] in the external validation dataset.

GLUCOSE policy analysis

We further assessed the model by evaluating the time in range (TIR) of 70–180 mg/dL for glucose level when the actual clinician administered insulin dose was similar or different from the dose recommended by GLUCOSE. In the internal testing set, 27.3% of the time patients received insulin doses from clinicians identical to those recommended by GLUCOSE, while in the external validation dataset, this occurred 20.3% of the time. As shown in Fig. 2b, patients who received insulin doses like those suggested by GLUCOSE had the highest average TIR in both the internal testing set and external validation dataset. The TIR decreased as the difference of model recommended doses minus clinician administered doses increased, indicating that the model identifies areas for improvement in insulin management. For example, at more negative cumulative differences, where the average glucose is also lower, GLUCOSE suggests less insulin to mitigate the risk of hypoglycemia (Fig. 2b, c). Conversely, at more positive cumulative differences, where average blood glucose is higher and average

Table 1 | Patient characteristics

		Development cohort	External validation cohort	P value
Admission Age, median [Q1, Q3]		68.0 [61.0,76.0]	67.0 [60.0,75.0]	0.093
Gender, n (%)	Female	1746 (28.4)	213 (32.8)	0.02
	Male	4402 (71.6)	436 (67.2)	
Height, median [Q1,Q3]		173.0 [165.0,178.0]	170.2 [162.6,177.8]	0.059
Weight, median [Q1,Q3]		83.2 [72.0,96.0]	85.6 [72.6,99.7]	0.005
BMI, median [Q1,Q3]		28.2 [25.1,32.1]	29.2 [25.8,33.9]	<0.001
24 h Sequential Organ Failure Assessment Score, median [Q1,Q3]		5.0 [3.0,7.0]	6.0 [4.0,8.0]	<0.001
Race, n (%)	Asian	136 (2.2)	15 (2.3)	<0.001
	Black	222 (3.6)	51 (7.9)	
	Hispanic	175 (2.8)	79 (12.2)	
	Native American	11 (0.2)		
	Other	219 (3.6)		
	Unknown	841 (13.7)	72 (11.1)	
	White	4544 (73.9)	432 (66.6)	
Coronary Artery Bypass Graft, n (%)		4390 (71.4)	516 (79.5)	<0.001
Valve Repair/Replacement, n (%)	Aortic	835 (13.6)	125 (19.3)	<0.001
	Pulmonary	0 (0.0)	0 (0.0)	<0.001
	Mitral	332 (5.4)	69 (10.6)	<0.001
	Tricuspid	221 (3.6)	1 (0.2)	<0.001
	Unspecified	1424 (23.2)	11 (1.7)	<0.001
Type 1 Diabetes, n (%)		33 (0.5)	2 (0.3)	0.77
Type 2 Diabetes, n (%)		868 (14.1)	59 (9.1)	<0.001
End Stage Renal Disease, n (%)		131 (2.1)	17 (2.6)	0.503
Congestive Heart Failure, n (%)		1665 (27.1)	52 (8.0)	<0.001
Chronic Obstructive Pulmonary Disease, n (%)		297 (4.8)	65 (10.0)	<0.001
History of Hypertension, n (%)		3814 (62.0)	177 (27.3)	<0.001
History of Myocardial Infarction, n (%)		1782 (29.0)	56 (8.6)	<0.001

Percentages may exceed 100% for types of surgery as patients may undergo both coronary artery bypass grafting and valvular surgery, including multiple valves, simultaneously.

TIR is worse, the model suggests higher insulin doses to avoid hyperglycemia. To explore subgroup-specific performance, we conducted a TIR analysis stratified by sex, race, and diabetic status (Supplementary Fig. 1). Across all subgroups, GLUCOSE achieved the highest average TIR when its recommended insulin dose matched the clinician-administered dose. This consistent pattern across all groups suggests that GLUCOSE performs well across these subgroups of patient populations.

The action distribution of Fig. 2c further illustrates these dosing patterns. Across glucose ranges below 180 mg/dL, GLUCOSE consistently recommends lower average insulin doses than clinicians, reflecting a guideline-aligned strategy that prioritizes maintaining glucose between 140 and 180 mg/dL while reducing the risk for hypoglycemia. The recommended insulin dose starts to increase after this threshold and surpasses the clinician doses when glucose levels were above 200 mg/dL, demonstrating a proactive approach to correcting significant hyperglycemia aligning with the recommendations to avoid hyperglycemia while minimizing the risk of hypoglycemia.

To characterize clinician-model disagreement, we analyzed the clinical features associated with the top and bottom deciles of absolute differences between clinician-administered insulin and GLUCOSE-predicted insulin, corresponding to the highest and lowest disagreement, respectively. Across both internal and external cohorts, the largest disagreements occurred near glucose values of approximately 140 mg/dL. This range lies near the lower boundary of the 140–180 mg/dL target recommended for glucose management among critically ill patients. In the internal testing set, clinicians administered an average of 5.9 units of insulin in high-disagreement cases

compared to the model's 1.8 units, with a mean glucose of 142 mg/dL. Similarly, in the external validation set, clinicians gave 6.7 units versus the model's 1.6 units at a mean glucose of 139.5 mg/dL.

To gain insight into model representations and ensure its clinical interpretability, we derived feature importances for GLUCOSE using SHapley Additive exPlanations (SHAP) (Supplementary Fig. 2)³³. This analysis revealed that the most heavily weighted features align well with clinical intuition. Notably, recent and historical glucose measurements emerged as key predictors, underscoring the value of capturing real-time trends. Additionally, indicators of patient acuity, which may influence stress-induced hyperglycemia, such as the use of and duration of mechanical ventilation, Sequential Organ Failure Assessment (SOFA) score, Elixhauser Comorbidity Index, and the type of surgery, were weighed heavily. These findings suggest that GLUCOSE uses clinically relevant information in its decision making.

Human evaluations of GLUCOSE

For clinical applicability and robustness, we conducted a multi-phased human evaluation. In the first phase, two senior endocrinologists, each with over 10 years of clinical experience, provided their recommendations for hourly insulin dosing for the first day after cardiac surgery for 10 patients in both internal testing and external validation datasets. To allow the endocrinologists to provide the most accurate dosing schemes to use as a reference, we provided them with the entire time series of patient data, including the insulin doses actually administered by the treating clinicians, and resultant glucose levels. We compared the hourly insulin doses

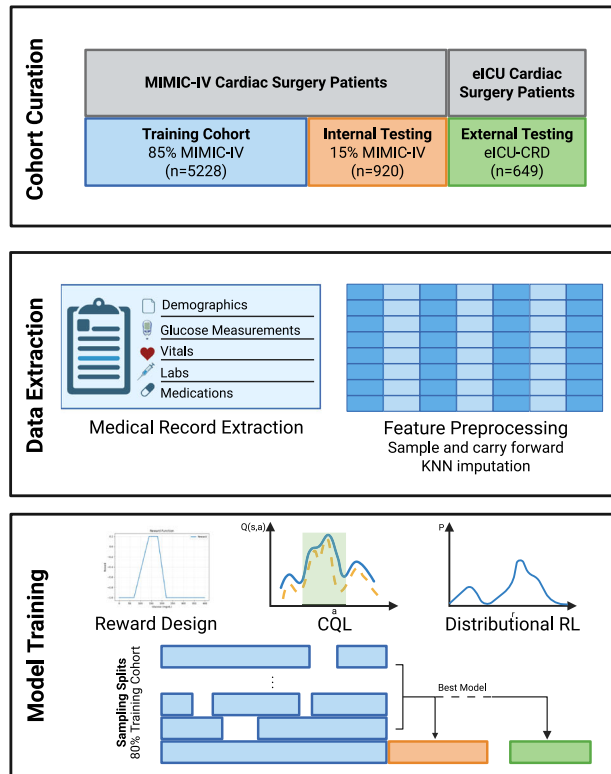
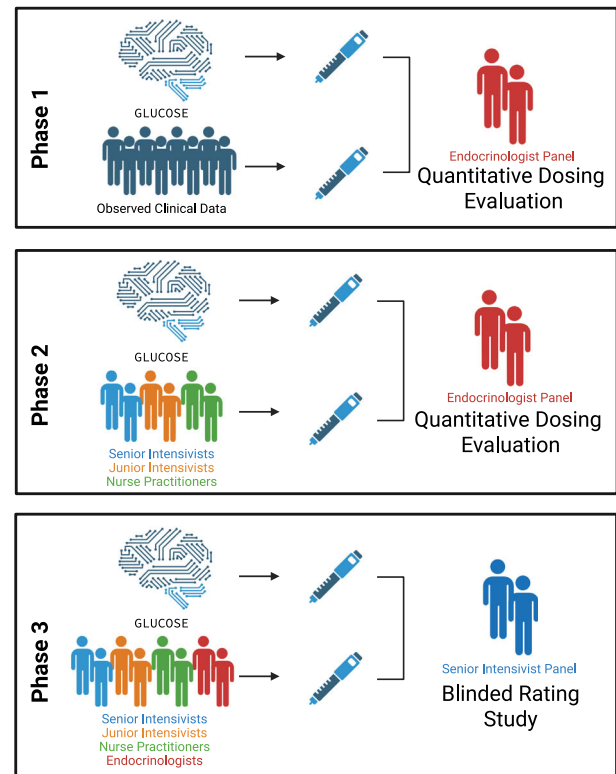
a. Computational Modeling**b. Human Evaluation**

Fig. 1 | Study overview. a Schema for model development, testing, and selection. **b** Overview of clinician validation study. Created in BioRender. Desman, J. (2025) <https://BioRender.com/k11i184>.

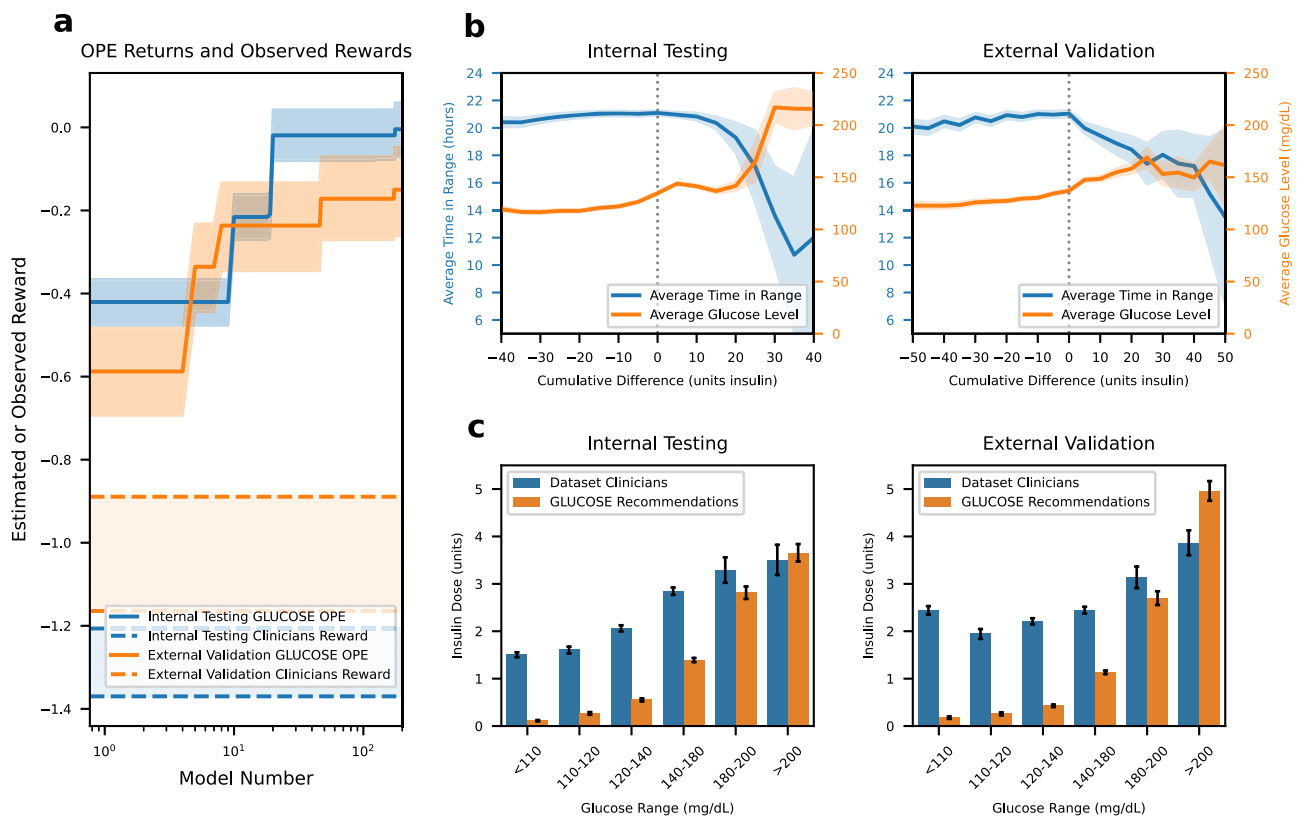


Fig. 2 | GLUCOSE performance. a OPE counterfactual estimated performance of the model computed by FQE (solid lines) compared to the returns by the treating clinicians (dotted lines) with 95% CI. **b** Comparison of TIR and average glucose

relative to insulin dosing differences with 95% CI. **c** Average insulin doses across several glucose ranges with 95% CI.

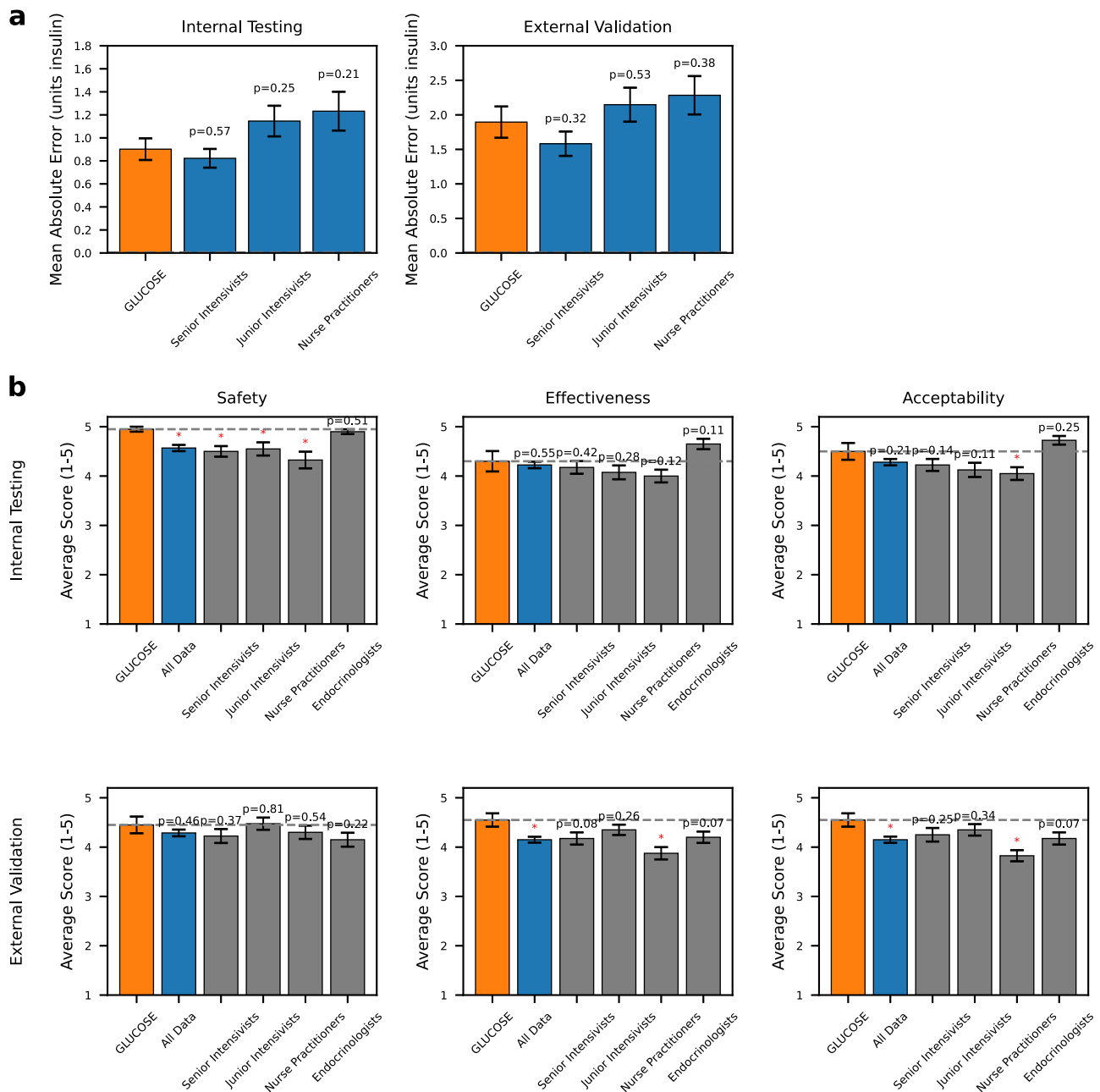


Fig. 3 | Human validation study results. a MAE of clinician groupings and GLUCOSE relative to an endocrinologist baseline with standard error of the mean. **b** Blinded ratings across safety, effectiveness, and acceptability of all clinicians by a blinded senior intensivist panel with standard error of the mean.

recommended by GLUCOSE, which unlike the endocrinologists had access only to the current state, to those actually administered by clinicians using the average endocrinologist doses as the reference. Across both datasets, GLUCOSE achieved significantly lower mean absolute error (MAE) in hourly insulin dosing, indicating that its dosing scheme more closely aligned with the recommendations of the endocrinologists. In the internal testing set, GLUCOSE had an MAE of 0.9 units compared to the treating clinician's 1.97 units MAE ($p < 0.001$). In the external validation dataset, GLUCOSE had an MAE of 1.90 units compared to the treating clinician's MAE of 2.24 units ($p = 0.003$).

In the second phase, two senior cardiac intensivists (> 5 years' experience), two junior cardiac intensivists (< 5 years' experience), and two cardiac intensive care unit nurse practitioners provided their recommendations for hourly insulin doses for the same patients. These clinicians were also provided with the entire time series of data, actual insulin administration record, and glucose levels to allow them to generate their most

retrospectively optimal possible human policies. We then compared the GLUCOSE recommended doses, which again only had access to a single state of information at the current timestep, to those recommended by these 6 clinicians, with the endocrinologist recommendations as the reference (Fig. 3a). In internal testing, GLUCOSE achieved an MAE of 0.90 units compared to that of senior intensivists' 0.82 unit MAE ($p = 0.57$), junior intensivists' 1.15 unit MAE ($p = 0.25$), and nurse practitioners' 1.23 unit MAE ($p = 0.21$). In external validation, GLUCOSE achieved an MAE of 1.90 units compared to that of senior intensivists' 1.58 units ($p = 0.32$), junior intensivists' 2.15 units ($p = 0.53$), and nurse practitioners' 2.28 units ($p = 0.38$). Although the differences in MAE did not reach statistical significance, GLUCOSE demonstrates a trend toward lower MAEs than that of junior intensivists and nurse practitioners when compared against endocrinologists as the reference.

In the final phase, we conducted a blinded evaluation of GLUCOSE and all 8 clinician dosing recommendations using an expert panel of 2 separate

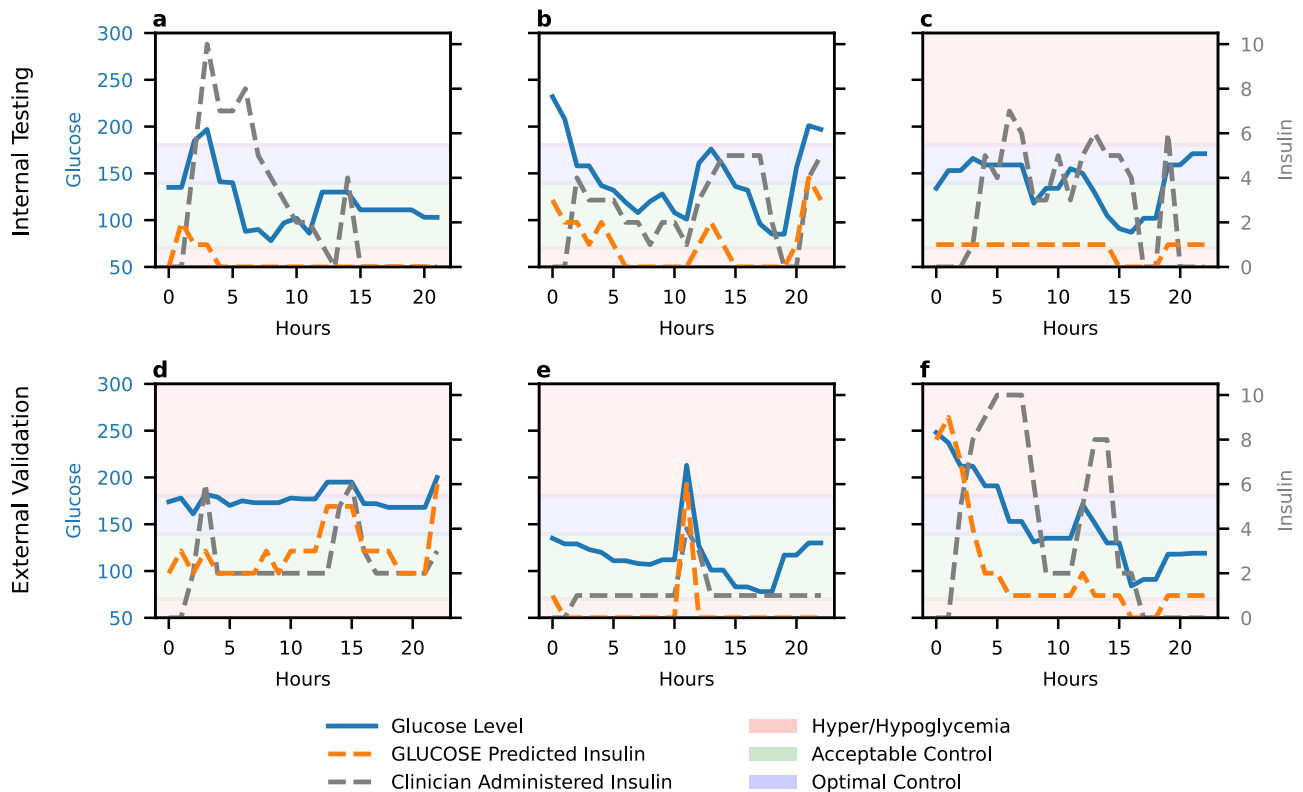


Fig. 4 | Case examples. Representative case examples to compare GLUCOSE's insulin dosing recommendations compared with actual clinician administered insulin in internal testing (a–c) and external validation (d–f) patients. Lines indicate glucose levels and insulin doses. Colored bands indicate glycemic ranges.

senior intensivists to assess practical safety, effectiveness, and acceptability of the model's recommended insulin doses. The two additional senior intensivists used a 5-point Likert scale to assess the safety (to reduce hypoglycemia), effectiveness (if the regimen would bring glucose into an acceptable range), and acceptability (if the regimen would be acceptable in a clinical scenario) of each recommended insulin regimen for the same group of 20 patients. In both internal testing and external validation datasets, GLUCOSE's rated safety, effectiveness, and acceptability demonstrated either comparable performance or statistically significant improvements over all human policies (Fig. 3b). Notably, GLUCOSE performed at or above the level of senior cardiac surgery intensivists across all domains in both the internal testing and external validation datasets. This demonstrates GLUCOSE's reliability and consistent high-level performance across diverse clinical scenarios.

To illustrate GLUCOSE's real time decision making, we provide representative case examples comparing its insulin dosing recommendations with actual clinician-administered doses (Fig. 4). Overall, GLUCOSE consistently demonstrates dynamic and personalized insulin dosing strategies, adapting to changes in glucose trajectories. Across randomly selected internal testing and external validation patients, GLUCOSE provided timely insulin adjustments, often moderating dosing to avoid overshooting glucose targets. These cases highlight how GLUCOSE responds to evolving patient conditions and targets an optimal glucose range more in line with STS guidelines.

Evaluation of GLUCOSE's recommendations among excluded patients

Finally, we evaluated GLUCOSE's recommendations in subsets of the external validation dataset that were excluded from the primary analysis due to the presence of ambiguous administration of insulin, vasopressors or inotropes. These patients had documented insulin, vasopressor, or inotrope administration but with insufficient information to determine the exact timing or dose – an issue commonly reported in the multicenter eICU-CRD database³⁴. Based on the affected medication, we performed this evaluation separately in

patients with only ambiguous insulin data (3,001 patients) and in those with ambiguous data for both insulin and vasopressors/inotropes (1,804 patients). As there were only 33 patients with non-ambiguous insulin but ambiguous vasopressor/inotrope data, we excluded them from this analysis.

The external validation cohort, the ambiguous insulin subset, and the ambiguous medication subset were similar in terms of age, gender, and weight. However, these additional subsets included a higher proportion of white patients (66.6% vs 81.1% vs 89.7%, $p < 0.001$) and fewer patients with type 2 diabetes (9.1% vs. 3.1% vs. 1.2%, $p < 0.001$). While there were statistically significant differences in the average glucose levels (134.5 mg/dL vs. 132.2 mg/dL vs. 130.6 mg/dL, $p < 0.001$) these small differences are not clinically meaningful. Full demographic analysis can be found in Supplementary Table 2.

Due to the lack of accurately recorded insulin in these subgroups, we were unable to perform OPE or direct comparisons which depend on accurate insulin administration records. However, the overall distribution of model recommended actions was comparable across datasets (Supplementary Fig. 3). Although there were statistically significant differences, all differences in average insulin across all glucose ranges were less than half a unit and therefore not clinically significant (Supplementary Fig. 3).

Discussion

In this study we have developed GLUCOSE, a distributional offline RL based algorithm, that dynamically suggests personalized regular insulin dosing for patients in the first day after cardiac surgery. The algorithm was validated on an independent multicenter dataset and further demonstrated its robustness and safety through rigorous human evaluations.

Hyperglycemia early after cardiac surgery is associated with higher rates of post-operative infections^{4,5}, acute kidney injury^{3,6–8}, cardiac arrhythmias³, longer length of stay³, and higher mortality^{6–9}. This underscores the importance of glucose control in the early post-operative period. Moreover, research indicates that the harmful effects of hyperglycemia are dose-dependent, with longer exposure and higher glucose levels leading to

worse outcomes³⁵. Therefore, it is essential to manage both the severity and duration of hyperglycemia. The STS recommends maintaining blood glucose levels below 180 mg/dL after cardiac surgery¹⁰. To achieve this target, most cardiac surgery centers employ institutional protocols for managing hyperglycemia^{15,36–38}. However, a significant challenge in early postoperative glucose management is that insulin, the primary treatment for hyperglycemia, has a narrow therapeutic window³⁹. Since these protocolized regimens often fail to account for individual patient variability in real-world settings^{13, 14}, hypoglycemia becomes a significant risk, particularly with intensive insulin dosing schemes^{40,41}. Hypoglycemia, defined as a blood glucose level <70 mg/dL, can trigger increased sympathetic activity leading to increased heart rate or arrhythmias⁴², impairment of autonomic cardiac reflexes⁴³, poor neurological outcomes⁴⁴, and death¹². This hypoglycemia is seen in 5–21% patients after cardiac surgery^{11,12,40} prompting a more conservative insulin dosing which, in turn, can result in persistent hyperglycemia. Thus, not surprisingly, these protocols frequently fall short, with only 15% of patients reaching the recommended glucose levels without hypoglycemia on the first day after surgery¹¹, which is the most critical and dynamic period after cardiac surgery.

Clinicians review over 1300 data points per patient each day, making it difficult to effectively use all this information for clinical decision-making⁴⁵. An algorithm that can systematically process and interpret these data points can significantly enhance clinician workflow while improving patient outcomes. The GLUCOSE model addresses this by evaluating over 70 features, such as vasopressor doses, SOFA score, mechanical ventilation needs, past glucose values, and BMI, every hour. It recommends personalized insulin doses that account for the patient's evolving clinical status. Importantly, the algorithm prioritizes features that are clinically relevant, as reflected in the feature importance analysis. Consistently, the GLUCOSE dosing scheme outperforms traditional clinician-driven dosing strategies in terms of estimated average performance.

The TIR for glucose was highest when the administered insulin dose closely matched the model's recommendations. As the discrepancy between the administered doses and GLUCOSE's suggested doses increased, the time in range decreased. Notably, when the difference was negative, meaning GLUCOSE recommended less insulin than what was administered, the average glucose level was lower. Conversely, with positive differences—where GLUCOSE suggested more insulin than what was given—the average glucose level was higher. This indicates that aligning insulin doses more closely with GLUCOSE's recommendations could potentially increase time in range and reduce glucose variability, which is associated with worse clinical outcomes^{13,14}.

Although there have been algorithms developed to assist clinicians in insulin doses, they are mostly limited to simulated settings without any human evaluations, include exclusively patients with diabetes, and none specifically target post-cardiac surgery patients^{46–48}. To ensure clinical applicability and acceptability of our study we performed a comprehensive 3-phase human evaluation inspired by prior work⁴⁶, which is a significant strength of our study. The results of our multi-phased human evaluation underscore the clinical robustness and reliability of the GLUCOSE algorithm in guiding insulin dosing for post-cardiac surgery patients. The significant reduction in MAE achieved by GLUCOSE compared to observed clinician dosing across both internal and external datasets highlights the algorithm's agreement with rigorous clinical evaluation. Particularly noteworthy is GLUCOSE's performance in the final phase of the evaluation, where it was assessed by senior intensivists on safety, effectiveness, and acceptability. The algorithm was either comparable to or exceeded the standards set by experienced clinicians, including senior cardiac surgery intensivists. It is important to note, that unlike GLUCOSE, which only had access to patient data till each current time-step, the clinicians that performed human evaluations had access to the entire patient time series of data. This made their approach nearly optimal, against which GLUCOSE's performance was measured. In reality, clinicians also only have access to data up to the current time step, making GLUCOSE's performance particularly notable in this context. This suggests that GLUCOSE can provide a

valuable tool in the management of hyperglycemia in this critically ill patient population, offering a level of reliability and clinical applicability that is on par with traditional human-driven dosing strategies. The ability of GLUCOSE to maintain high performance across diverse clinical scenarios further supports its potential integration into clinical practice, where it could enhance patient outcomes by reducing variability in insulin dosing and minimizing the risks associated with both hyperglycemia and hypoglycemia.

Incorporation of distributional RL is another significant strength of this study. Even among patients with seemingly similar clinical profiles, there can be considerable variation in physiological responses. Distributional RL is particularly well-suited to address this challenge, as it quantifies the intrinsic uncertainty within a Markov Decision Process (MDP), which is characteristic of stochastic environments²³. By learning to approximate the distribution of potential outcomes, this approach strengthens the model by preparing it to handle the inherent uncertainties of real-world clinical settings.

Given the variability in hospital protocols and patient populations, the successful integration of GLUCOSE into clinical practice may require site or unit specific customization. For example, finetuning the model for distinct clinical scenarios, such as managing sepsis in the ICU or treating patients on the wards with subcutaneous insulin, could broaden its applicability beyond the post-cardiac surgery context. Future work should explore the use of transfer learning to adapt GLUCOSE for the general wards, non-cardiac ICUs, or other settings characterized by unique nutritional and metabolic demands. Such tailored adaptations would not only enhance the model's generalizability but also promote the widespread use of dynamic insulin titration protocols, ultimately improving patient outcomes by reducing dosing variability and mitigating hyper- and hypoglycemia risk. To further promote generalizability, we limited GLUCOSE's input features to routinely collected ICU data that are standardized across institutions and consistently available in electronic health records. This design choice allows GLUCOSE to operate effectively across heterogeneous hospital systems, such as those included in the eICU-CRD dataset. GLUCOSE's strong performance in this multicenter external validation cohort, including under scrutiny of senior clinicians, supports the robustness of the overall approach.

We envision GLUCOSE as a clinical decision support tool integrated into electronic health record systems to provide real-time, personalized insulin dosing recommendations. GLUCOSE is designed to integrate seamlessly into existing ICU workflows. The model utilizes routinely collected clinical data, ensuring that no additional data collection burden is placed on healthcare providers. It can be deployed within existing electronic health record systems with minimal technical adjustments, making it accessible to a wide range of hospitals. By continuously analyzing patient data, GLUCOSE moves beyond standardized protocols to deliver tailored insulin management that adapts to each patient's evolving clinical condition without disrupting patient care processes. This would, however, require an initial silent deployment to evaluate its performance against current practices, followed by a prospective clinical trial to rigorously assess its safety and efficacy.

Successful real-world deployment will also require addressing key regulatory and operational considerations. These include ensuring compliance with institutional policies and federal privacy regulations, such as the Health Insurance Portability and Accountability Act (HIPAA), robust protocols for ethical oversight and patient safety, and overcoming technical challenges of integrating the system into existing electronic health record platforms. Successful navigation of these hurdles will require close collaboration with hospital IT departments, clinical leadership, and institutional stakeholders, as well as sustained investment in implementation infrastructure.

While GLUCOSE demonstrates significant potential, several limitations should be considered when interpreting these results. First, although our retrospective study strongly supports the use of GLUCOSE as a clinical decision support tool, these findings require validation through prospective studies and clinical trials involving large and diverse patient cohorts. Second,

the GLUCOSE model has been trained, tested, and externally validated only for the first day following cardiac surgery. Though this is the most dynamic time-period for patients after cardiac surgery, expanding this work to evaluate insulin regimens over longer postoperative periods would further enhance GLUCOSE's clinical utility. Future versions of GLUCOSE could extend the observation window beyond the first 24 h or leverage transfer learning to enable adaptation to longer-term glucose management. Third, our current algorithm does not incorporate explicit dietary data. While nutrition is a known contributor to glycemic variability, there are several considerations that mitigate its impact in our current context. During the initial 24 h following cardiac surgery, patients typically receive minimal or no oral intake due to postoperative recovery protocols, mechanical ventilation, and anesthesia. This substantially reduces the influence of nutritional intake on glucose dynamics during this early window, which is the focus of our current study. Additionally, in real-world clinical settings, precise and time-stamped dietary data are rarely collected as part of routine care, making consistent integration into algorithmic models challenging. The need for such granular data could also hinder scalability and clinical adoption. Notably, prior studies have demonstrated that RL algorithms can achieve effective glycemic control without explicit meal information⁴⁷. This is likely because postprandial glucose fluctuations are captured within the glucose time series itself, allowing the model to learn latent representations of meal-related effects. By implicitly capturing the impact of nutrition through glucose dynamics, GLUCOSE reduces dependence on non-standard data inputs while maintaining clinical relevance. Nonetheless, we acknowledge that the exclusion of dietary data may limit performance in scenarios where nutritional intake becomes more variable, such as during extended ICU stays or in general ward settings. We view this as an important area for future exploration as we move toward broader deployment of the model. Finally, to ensure accurate training and validation of the model, we restricted ourselves to patients that had accurately documented doses of medications such as insulin, given that it was the action, and vasopressors/inotropes, which indicate the risk of disease severity and thus may portend a higher risk of hyperglycemia, in both development and external validation cohorts. With this, we did not need to exclude any patients in MIMIC-IV, but had to exclude 4,838 patients in eICU-CRD dataset. This missingness in eICU-CRD is a known issue with the dataset³⁴, but with patients from over 200 hospitals, it is a highly heterogeneous dataset and thus remains impactful for external validation. To ensure that our model's recommendations still generalize appropriately in the excluded subset of the external validation dataset we assessed the distribution of recommendations of the model in the subset with just ambiguous insulin data, and in the subset with ambiguous data about both insulin and vasopressor/inotropes. We found that the distribution of recommendations was very similar in the 3 groups, with no clinically meaningful differences in actions, which suggests good generalizability of the model.

In summary, we have developed and externally validated GLUCOSE, a distributional RL based model to dynamically optimize glucose management in cardiac surgery patients. The comprehensive three-phase human evaluations support GLUCOSE's clinical robustness and safety, demonstrating its effectiveness in real-world settings and its performance on par with or surpassing that of experienced clinicians. Future studies should be focused on randomized controlled trials to further evaluate the effectiveness and safety of GLUCOSE in diverse clinical settings.

Methods

Study design and databases

For this retrospective study, we used the MIMIC-IV database²⁵ to develop the GLUCOSE algorithm (Development dataset). MIMIC-IV is a single-center database constructed from deidentified ICU admissions at the Beth Israel Deaconess Medical Center from 2008 to 2019. We externally validated the derived policy using the heterogeneous eICU Collaborative Research Database (eICU-CRD)²⁶ (External Validation Dataset). eICU-CRD is constructed from over 200,000 de-identified admissions to 208 United States hospitals between 2014 and 2015.

Study population

We included all adult patients (age ≥ 18 years) who were admitted to ICU after cardiac surgery. We used ICD-9-PCS and ICD-10-PCS codes to identify patients who underwent cardiac surgery in MIMIC-IV database (Supplementary Table 3). The eICU-CRD database does not include ICD-9-PCS or ICD-10-PCS procedure codes. As per prior literature²⁶, we have identified patients admitted to the ICU after cardiac surgery using the "admissiondx" table that provides the primary diagnosis for ICU admissions (Supplementary Table 4). We excluded patients who died within first 24 h of ICU admission, had ambiguous medication administration information such that it did not allow us to calculate the exact dose of medication administered, or did not have available glucose levels within first three hours of documented ICU admission time after surgery. As our focus was to develop a policy to personalize the administration of regular insulin, we excluded patients who received other short acting insulins (aspart, lispro, NPH, insulin 70/30) (Supplementary Fig. 4).

Feature extraction and preprocessing

We extracted information about patient demographics (age, sex, race), comorbidities (history of diabetes, hypertension, end stage renal disease, chronic obstructive pulmonary disease, asthma, prior myocardial infarction, congestive heart failure, Elixhauser comorbidity score), laboratory values (complete blood count, comprehensive metabolic panel, coagulation studies, and blood gases), vital signs (systolic blood pressure, diastolic blood pressure, mean arterial pressure, heart rate, respiratory rate, temperature, and oxygen saturation), vasopressor and inotrope doses, mechanical ventilation status, and SOFA scores. We extracted the data as multidimensional discrete time series in 1-h time intervals, with features summed or averaged as clinically appropriate. We excluded features with over 30% missingness. In line with standard approach to handling missingness in these data, we used forward fill imputation for all features with k -nearest neighbor (k -NN) imputation ($k = 5$) to impute any remaining missing data^{28,49}. Only the first 24 h of data for each patient was utilized. All features were checked for outliers using a frequency histogram and descriptive statistics. Errors were corrected as appropriate, such as conversion of temperature to Fahrenheit to Celsius. The full feature list can be found in the Supplementary Table 1. All features across all datasets were normalized into range $[0, 1]$ based on the training set to improve training stability.

Our outcome was appropriate glucose control, defined as an hourly glucose level between 70–180 mg/dL^{10,50}, in the first day after cardiac surgery. Consequently, we began recording timesteps from the availability of the first glucose level measurement after admission to ICU.

Computational modeling

We used conservative Q learning (CQL), a state-of-the-art offline RL algorithm that allows model to suggest clinical actions while regularizing the learned policy to mitigate overestimation in low-coverage or out-of-distribution state-action pairs⁵¹. CQL was chosen over other offline RL methods, such as Batch-Constrained Q-Learning (BCQ), Behavior Regularized Actor Critic (BRAC), and Twin Delayed Deep Deterministic Policy Gradient with Behavior Cloning (TD3 + BC), because it explicitly and conservatively regularizes the learned policy by penalizing actions outside the dataset distribution, while still allowing for strategic generalization^{27,52,53}. CQL is considered among the state-of-the-art in offline RL due to its strong performance across standard benchmarks and its robust handling of out-of-distribution actions^{51,54}. Its conservative approach is particularly suited to our domain, where insulin management involves high-stakes decisions and a narrow therapeutic index. Therefore, CQL helps ensure that the policy remains grounded in safe and high-reward actions observed in the data, mitigating the risks associated with extrapolating to unsupported state-action pairs. To further enhance the model's understanding of uncertainty and risk, we integrated CQL with distributional RL, an approach that characterizes the entire distribution of potential outcomes rather than focusing only on the expected reward. By capturing the full range of possible patient

responses, this methodology enables more nuanced decision-making, particularly for rare but critical events such as hypoglycemia, which traditional RL methods may underestimate. All models use a multi-layer perceptron (MLP) network with 3 512-dimension hidden layers. This integration is crucial for making safe and effective decisions especially with clinical actions that have a narrow therapeutic index, such as insulin dosing. To achieve this, we incorporated Implicit Quantile Networks (IQN) into CQL, leveraging the strengths of distributional RL to better model the variability in patient responses, thereby improving the robustness of GLUCOSE²⁴. Unlike other distributional methods that require explicitly defining the number of quantiles, IQN adds a layer which flexibly learns the full return distribution by sampling from continuous quantile values during training, allowing it to approximate the entire outcome distribution without fixed bins. This enables a more comprehensive representation of potential outcomes while improving upon its non-distributional counterparts^{22–24}. As a result, the model can better capture clinical uncertainty and make decisions around nuanced risk profiles, particularly in settings with high variability. To the best of our knowledge, this is the first application of integration of CQL with distributional Q functions in healthcare.

Finally, we implemented a batch training sampling strategy for offline RL, which avoids overregularization by low-return actions, allowing the learned policy to reflect more high-return trajectories⁵⁵.

State space

RL typically considers problems as MDPs. An MDP can be represented as a tuple of (s_t, a, r, s_{t+1}) for each time step t . Here, s_t represents a vector observation of features at that hour index t , and s_{t+1} represents the state at the next hour index after taking action a . The reward, r , is given for taking action a at state s_t .

We used the features derived from demographics, comorbidities, laboratory values, vital signs, medications, mechanical ventilation, and SOFA scores binned into hourly time-steps to develop the state space. Based on previous literature, we incorporated the prior four hours of glucose values, when available, into the RL model⁵⁶. To provide additional context, we included information on glucose level changes during this period and calculated the ratio of glucose change to insulin dose for each hour, with the minimum insulin dose set at 0.1 for this calculation.

Action space

In our offline RL model, actions are defined as the amount of regular insulin administered each hour, utilizing a continuous action space. For ease of interpretability, we have rounded the recommended insulin doses to the nearest integer. This practice aligns the model's recommendations with practical clinical standards and facilitates the clinical implementation of suggested doses. Additionally, we capped the insulin doses recommended and observed by GLUCOSE at a maximum of 10 units per hour. This decision was based on both data-driven considerations and clinical safety parameters. The mean hourly insulin dose in MIMIC-IV was 2.2 ± 3.0 units/hour, with doses exceeding 10 units/hour accounting for only 2.29% of all administered doses. From a clinical perspective, given insulin's narrow therapeutic index, this cap serves as a safeguard to minimize the risk of hypoglycemia. It also reinforces the importance of maintaining a physician-in-the-loop framework for cases that may necessitate higher doses, ensuring safety and clinician oversight.

Reward

We designed our reward to maximize optimal glycemic control while strongly discouraging behavior that would result in both hypoglycemia (glucose <70 mg/dL) and hyperglycemia (glucose >180 mg/dL). We provided a maximum reward of +0.2 within the 140–180 mg/dL range and penalties become increasingly negative, down to -1, outside that range. We chose a reward with relatively low magnitude to improve training stability, and a relatively negative reward to disincentive any out of range values⁵⁷. The

reward is outlined in Eq.(1):

$$r = \begin{cases} -1, & x < 70 \\ 3x/175 - 2.2, & 70 \leq x < 140 \\ 0.2, & 140 \leq x < 180 \\ -0.03x + 5.6, & 180 \leq x < 220 \\ -1, & x \geq 220 \end{cases} \quad (1)$$

To ensure the safety of insulin dosing recommendations, considering insulin's narrow therapeutic window, we implemented an exponentially increasing penalty that serves to discourage large overcorrections and promotes more cautious dosing adjustments⁵⁸:

$$r_t = r_t - (0.001a_t^2) \quad (2)$$

GLUCOSE model training and external validation

To train the policies, we first split the development dataset into 85% training and 15% internal testing sets. Since RL is particularly subject to high stochastic training variation⁵⁷, we sought to mitigate sampling and stochastic biases by training multiple models on subsampled 80% splits of the training set. Each training run and sampling split utilized a unique seed to ensure different training sets and distinct sampling order while maintaining reproducibility. We continued training models until substantial and significant improvements over the clinician policies were observed in the 15% internal testing set. The final model (GLUCOSE) was selected as the one that had the highest lower bound of the 95% CI for estimated performance returns in the internal testing set²⁸. We then evaluate GLUCOSE on the external validation set (Fig. 1).

Training was conducted in batches of 256, with actor and critic learning rates of $1e^{-4}$ and $3e^{-4}$, respectively. The discount factor γ was set at 0.67, corresponding to a 3 h effective horizon (calculated as $1/(1-\gamma)$). Discount factors are problem specific, and the choice of a lower discount factor is critical in the context of this problem. Higher discount factors, such as those exceeding 0.95, extend the effective horizon beyond the episode length, resulting in future rewards being weighted nearly as heavily as immediate rewards. Glucose levels can change rapidly, which could lead to suboptimal policy development as the model may either overly prioritize distant future rewards unaffected by the current state or become insensitive to immediate low reward states. Using a lower discount factor aligns the temporal focus of the model to ensure it remains responsive to rapidly changing glucose levels. A dropout rate of $p = 0.1$ was applied during training to improve policy generalization. All models were implemented in Python 3.8.2 using d3rlpy⁵⁹.

Off-policy evaluation

We used fitted-Q-evaluation (FQE) for offline policy estimation (OPE) to estimate the performance of the learned policies^{31,32}. FQE is effective in handling large policy deviations from observed behaviors as well as stochasticity, and it has shown consistency and calibration in various healthcare-specific benchmarks^{60,61}. Bootstrapping was applied across all episodes in the datasets to generate 95% confidence intervals by sampling with replacement 10,000 times. We then estimated the performance using FQE on both internal testing set and external validation dataset.

We further explored policy performance by analyzing the time in range (TIR) (70–180 mg/dL) in relation to the difference of GLUCOSE's dosing recommendations and clinician-administered doses (Fig. 2b). We calculated the cumulative differences as the model's predicted insulin doses minus the observed insulin doses over the first 24 h of ICU stay:

$$\Delta = \sum_t^T a_{t,RL} - a_{t,observed} \quad (3)$$

Cumulative differences are positive when the RL model recommends more insulin than what was administered, while negative differences

indicate that the model predicted a lower insulin dose compared to what was observed.

Estimation of feature importance

The interpretability of machine learning models is crucial in clinical care, where the rationale behind a model's predictions must be clear to ensure patient safety and informed decision-making. SHAP, a method grounded in cooperative game theory, assesses the contribution of each feature to a model's prediction by analyzing all possible combinations of feature values³³. In this study, we employ Permutation SHAP to estimate these contributions, as it provides a model-agnostic framework for elucidating model outputs.

Human evaluations

We further assessed the clinical validity of GLUCOSE in three separate phases of human evaluations. In the first phase, we compared the hourly insulin dosing recommended by GLUCOSE to that administered to patients in both internal testing and external validation datasets. Two senior endocrinologists (RS, AG), each with over ten years of clinical experience, provided their own hourly insulin dosing recommendations for 10 randomly selected patients in each cohort. Using the average hourly insulin doses recommended by endocrinologists as reference, we compared the insulin doses recommended by GLUCOSE to those administered to these patients.

In the second phase, we had two senior cardiac intensivists (> 5 years' experience) (PM, VN), two junior cardiac intensivists (< 5 years' experience) (DP, JG), and two cardiac intensive care unit nurse practitioners (AC, DG) provide their recommendations for hourly insulin doses in the first day after cardiac surgery for the same patients. We then compared the GLUCOSE recommended doses to those recommended by these clinicians, again using the average endocrinologist recommendations as references.

In the third phase, a panel of two senior intensivists (AS, GK) conducted a blinded evaluation of GLUCOSE against other clinician recommended insulin dosing schemes. Senior intensivists were selected for this phase because, in practice, these frontline clinicians are frequently responsible for making rapid decisions regarding glucose control in critically ill patients. They evaluated each dosing scheme for each patient using the following 5-point Likert scale questions:

1. Q1 (Safety)—How much risk for hypoglycemia does this regimen put the patient at? 1. Very high risk 2. High risk 3. Medium risk 4. Low risk 5. Minimal risk
2. Q2 (Effectiveness)—How effective is this regimen in bringing the glucose level within an acceptable range? 1. Not effective at all 2. Slightly effective 3. Moderately effective 4. Very effective 5. Extremely effective
3. Q3 (Acceptability)—How acceptable would this regimen be to you in clinical settings? 1. Strongly unacceptable 2. Unacceptable 3. Neutral 4. Acceptable 5. Strongly acceptable

Evaluation of model recommendations in excluded patient subsets

We also evaluated the GLUCOSE's recommendations using the part of external validation dataset that was excluded from the primary analysis due to presence of ambiguous insulin administration data, which prevented us from making direct comparisons or calculating OPE. Patients for which exact doses of insulin, vasopressors, or inotropes could not be resolved were separated and underwent the same exclusion criteria and preprocessing used for the primary external validation cohort. Any ambiguous data was zero-filled.

Statistical analysis

We performed comparisons of categorical features using Chi-squared test and continuous features using *t* test and Kruskal-Wallis test, as appropriate. All significance levels were set at $\alpha = 0.05$. To compare insulin doses administered by clinicians and GLUCOSE before hypo- and hyperglycemic episodes, we used Mann-Whitney U tests given the skewed distributions. To evaluate the accuracy of the insulin dosing schemes, we calculated the mean absolute error (MAE) between the insulin doses recommended by

various dosing schemes to those provided by endocrinologists. MAEs were determined by subtracting the endocrinologists' recommendations from the doses suggested by clinicians and the GLUCOSE system for each hourly dose administered or recommended for the 20 retrospectively reviewed patients. We then performed a *t* test to identify any significant differences in MAE between the observed clinicians' dosing and the endocrinologist's suggested dosing, and MAE between GLUCOSE's suggested dosing and the endocrinologist's suggested dosing. To assess differences in the average insulin doses across subsets of excluded patients, we used ANOVA tests for group-wise assessment and two sided *t*-tests for pairwise assessment. All statistical tests were conducted using Python 3.8.2 using SciPy⁶².

Data availability

All datasets used and analyzed in this present study are publicly available. MIMIC-IV and eICU-CRD data can be obtained via their online repositories at <https://physionet.org/content/mimiciv/2.2/> and <https://physionet.org/content/eicu-crd/2.0/>, respectively.

Code availability

The underlying code for this study is not publicly available for proprietary reasons. Code for GLUCOSE may be shared upon reasonable requests to the corresponding author.

Received: 10 December 2024; Accepted: 8 May 2025;

Published online: 27 May 2025

References

1. Najmali, S., Redford, D. & Larson, D. F. Hyperglycemia as an effect of cardiopulmonary bypass: intra-operative glucose management. *J. Extra Corpor. Technol.* **38**, 168–173 (2006).
2. Galindo, R. J., Fayman, M. & Umpierrez, G. E. Perioperative management of hyperglycemia and diabetes in cardiac surgery patients. *Endocrinol. Metab. Clin. North Am.* **47**, 203–222 (2018).
3. Moorthy, V., Sim, M. A., Liu, W., Chew, S. T. H. & Ti, L. K. Risk factors and impact of postoperative hyperglycemia in nondiabetic patients after cardiac surgery: A prospective study. *Med. (Baltim.)* **98**, e15911 (2019).
4. Lowden, E. et al. Evaluation of outcomes and complications in patients who experience Hypoglycemia after Cardiac Surgery. *Endocr. Pract.* **23**, 46–55 (2017).
5. Greco, G. et al. Diabetes and the association of postoperative hyperglycemia with clinical and economic outcomes in cardiac surgery. *Diab Care* **39**, 408–417 (2016).
6. Ruan, J. et al. Association between hyperglycemia at ICU admission and postoperative acute kidney injury in patients undergoing cardiac surgery: Analysis of the MIMIC-IV database. *J. Intensive Med.* **4**, 526–536 (2024).
7. Schmeltz, L. R. et al. Reduction of surgical mortality and morbidity in diabetic patients undergoing cardiac surgery with a combined intravenous and subcutaneous insulin glucose management strategy. *Diab Care* **30**, 823–828 (2007).
8. Ascione, R., Rogers, C. A., Rajakaruna, C. & Angelini, G. D. Inadequate blood glucose control is associated with in-hospital mortality and morbidity in diabetic and nondiabetic patients undergoing cardiac surgery. *Circulation* **118**, 113–123 (2008).
9. Giakoumidakis, K., Nenekidis, I. & Brokalaki, H. The correlation between peri-operative hyperglycemia and mortality in cardiac surgery patients: A systematic review. *Eur. J. Cardiovasc Nurs.* **11**, 105–113 (2012).
10. Lazar, H. L. et al. The Society of Thoracic Surgeons practice guideline series: Blood glucose management during adult cardiac surgery. *Ann. Thorac. Surg.* **87**, 663–669 (2009).
11. Williams, J. B. et al. Glycemic control in patients undergoing coronary artery bypass graft surgery: Clinical features, predictors, and outcomes. *J. Crit. Care* **42**, 328–333 (2017).

12. Johnston, L. E. et al. Postoperative Hypoglycemia Is Associated With Worse Outcomes After Cardiac Operations. *Ann. Thorac. Surg.* **103**, 526–532 (2017).
13. Martinez, M., Santamarina, J., Pavesi, A., Musso, C. & Umpierrez, G. E. Glycemic variability and cardiovascular disease in patients with type 2 diabetes. *BMJ Open Diabetes Res Care* **9**, e002032 (2021).
14. Rodbard, D. Glycemic variability: Measurement and utility in clinical medicine and research—one viewpoint. *Diab Technol. Ther.* **13**, 1077–1080 (2011).
15. Moghissi, E. S. et al. American Association of Clinical Endocrinologists and American Diabetes Association consensus statement on inpatient glycemic control. *Diab Care* **32**, 1119–1131 (2009).
16. Nguyen, M., Jankovic, I., Kalesinskas, L., Baiocchi, M. & Chen, J. H. Machine learning for initial insulin estimation in hospitalized patients. *J. Am. Med. Inf. Assoc.* **28**, 2212–2219 (2021).
17. Fitzgerald, O. et al. Incorporating real-world evidence into the development of patient blood glucose prediction algorithms for the ICU. *J. Am. Med. Inf. Assoc.* **28**, 1642–1650 (2021).
18. Zale, A. & Mathioudakis, N. Machine learning models for inpatient glucose prediction. *Curr. Diab Rep.* **22**, 353–364 (2022).
19. Montaser, E., Diez, J. L. & Bondia, J. Glucose prediction under variable-length time-stamped daily events: A seasonal stochastic local modeling framework. *Sensors (Basel)* **21**, 3188 (2021).
20. Kaelbling, L. P., Littman, M. L. & Moore, A. W. Reinforcement learning: A survey. *J. Artif. Intell. Res.* **4**, 237–285 (1996).
21. Levine, S., Kumar, A., Tucker, G. & Fu, J. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *arXiv* <https://doi.org/10.48550/arXiv.2005.01643> (2020).
22. Bellemare, M. G., Dabney, W. & Munos, R. in *International conference on machine learning*. 449–458 (PMLR).
23. Dabney, W., Rowland, M., Bellemare, M. & Munos, R. *Proceedings of the AAAI conference on artificial intelligence*.
24. Dabney, W., Ostrovski, G., Silver, D. & Munos, R. in *International conference on machine learning*. 1096–1105 (PMLR).
25. Johnson, A. E. W. et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci. Data* **10**, 1 (2023).
26. Pollard, T. J. et al. The eICU Collaborative Research Database, a freely available multi-center database for critical care research. *Sci. Data* **5**, 180178 (2018).
27. Fujimoto, S. & Gu, S. S. A minimalist approach to offline reinforcement learning. *Adv. neural Inf. Process. Syst.* **34**, 20132–20145 (2021).
28. Komorowski, M., Celi, L. A., Badawi, O., Gordon, A. C. & Faisal, A. A. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat. Med.* **24**, 1716–1720 (2018).
29. Peine, A. et al. Development and validation of a reinforcement learning algorithm to dynamically optimize mechanical ventilation in critical care. *npj Digital Med.* **4**, 32 (2021).
30. Lee, H. et al. Development and validation of a reinforcement learning model for ventilation control during emergence from general anesthesia. *npj Digital Med.* **6**, 145 (2023).
31. Le, H., Voloshin, C. & Yue, Y. in *International Conference on Machine Learning*. 3703–3712 (PMLR).
32. Paine, T. L., et al. Hyperparameter selection for offline reinforcement learning. *arXiv* <https://doi.org/10.48550/arXiv.2007.09055> (2020).
33. Lundberg, S. M. & Lee, S.-I. A unified approach to interpreting model predictions. *Advances in neural information processing systems* **30** (2017).
34. 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* <https://doi.org/10.1101/2023.04.20.23288845> (2023).
35. Navaratnarajah, M. et al. Effect of glycaemic control on complications following cardiac surgery: literature review. *J. Cardiothorac. Surg.* **13**, 10 (2018).
36. van den Berghe, G. et al. Intensive insulin therapy in critically ill patients. *N. Engl. J. Med.* **345**, 1359–1367 (2001).
37. Wiener, R. S., Wiener, D. C. & Larson, R. J. Benefits and risks of tight glucose control in critically ill adults: a meta-analysis. *Jama* **300**, 933–944 (2008).
38. Dellinger, R. P., et al. Surviving Sepsis Campaign: international guidelines for management of severe sepsis and septic shock: 2008. *Intensive Care Med.* **34**, 17–60 (2008).
39. Zaykov, A. N., Mayer, J. P. & DiMarchi, R. D. Pursuit of a perfect insulin. *Nat. Rev. Drug Discov.* **15**, 425–439 (2016).
40. Umpierrez, G. et al. Randomized controlled trial of intensive versus conservative glucose control in patients undergoing coronary artery bypass graft surgery: GLUCO-CABG trial. *Diab Care* **38**, 1665–1672 (2015).
41. Investigators, N.-S. S. Intensive versus conventional glucose control in critically ill patients. *N. Engl. J. Med.* **360**, 1283–1297 (2009).
42. Fisher, B. M., Gillen, G., Hepburn, D. A., Dargie, H. J. & Frier, B. M. Cardiac responses to acute insulin-induced hypoglycemia in humans. *Am. J. Physiol.* **258**, H1775–H1779 (1990).
43. Adler, G. K. et al. Antecedent hypoglycemia impairs autonomic cardiovascular function: implications for rigorous glycemic control. *Diabetes* **58**, 360–366 (2009).
44. Mohseni, S. Neurologic damage in hypoglycemia. *Handb. Clin. Neurol.* **126**, 513–532 (2014).
45. Manor-Shulman, O., Beyene, J., Frndova, H. & Parshuram, C. S. Quantifying the volume of documented clinical information in critical illness. *J. Crit. Care* **23**, 245–250 (2008).
46. Wang, G. et al. Optimized glycemic control of type 2 diabetes with reinforcement learning: a proof-of-concept trial. *Nat. Med.* **29**, 2633–2642 (2023).
47. Fox, I., Lee, J., Pop-Busui, R. & Wiens, J. in *Machine Learning for Healthcare Conference*. 508–536 (PMLR).
48. Zhu, T., Li, K., Herrero, P. & Georgiou, P. Basal glucose control in type 1 diabetes using deep reinforcement learning: An in silico validation. *IEEE J. Biomed. Health Inform.* **25**, 1223–1232 (2020).
49. Raghu, A., Komorowski, M., Celi, L. A., Szolovits, P. & Ghassemi, M. in *Machine Learning for Healthcare Conference*. 147–163 (PMLR).
50. Jacobi, J., et al. Guidelines for the use of an insulin infusion for the management of hyperglycemia in critically ill patients. *Crit. Care Med.* **40**, 3251–3276 (2012).
51. Kumar, A., Zhou, A., Tucker, G. & Levine, S. Conservative q-learning for offline reinforcement learning. *Adv. Neural Inf. Process. Syst.* **33**, 1179–1191 (2020).
52. Fujimoto, S., Meger, D. & Precup, D. in *International conference on machine learning*. 2052–2062 (PMLR).
53. Wu, Y., Tucker, G. & Nachum, O. Behavior regularized offline reinforcement learning. *arXiv* <https://doi.org/10.48550/arXiv.1911.11361> (2019).
54. Fu, J., Kumar, A., Nachum, O., Tucker, G. & Levine, S. D4rl: Datasets for deep data-driven reinforcement learning. *arXiv* <https://doi.org/10.48550/arXiv.2004.07219> (2020).
55. Hong, Z.-W., Agrawal, P., Combes, R. T. d. & Laroche, R. Harnessing mixed offline reinforcement learning datasets via trajectory weighting. (2023). <https://doi.org/10.48550/arXiv.2306.13085>
56. Emerson, H., Guy, M. & McConville, R. Offline reinforcement learning for safer blood glucose control in people with type 1 diabetes. *J. Biomed. Inform.* **142**, 104376 (2023).
57. Van Hasselt, H. P., Guez, A., Hessel, M., Mnih, V. & Silver, D. Learning values across many orders of magnitude. *Advances in neural information processing systems* **29** (2016).
58. Gu, W. & Wang, S. An improved strategy for blood glucose control using multi-step deep reinforcement learning. *arXiv* <https://doi.org/10.48550/arXiv.2403.07566> (2024).
59. Seno, T. & Imai, M. d3rlpy: An offline deep reinforcement learning library. *J. Mach. Learn. Res.* **23**, 1–20 (2022).

60. Voloshin, C., Le, H. M., Jiang, N. & Yue, Y. Empirical study of off-policy policy evaluation for reinforcement learning. *arXiv* <https://doi.org/10.48550/arXiv.1911.06854> (2019).
61. Tang, S. & Wiens, J. in *Machine Learning for Healthcare Conference*. 2–35 (PMLR).
62. Virtanen, P. et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nat. methods* **17**, 261–272 (2020).

Acknowledgements

This study was supported by NIH/NIDDK grant K08DK131286 (A.S.). This work was supported in part through the computational and data resources and staff expertise provided by Scientific Computing and Data at the Icahn School of Medicine at Mount Sinai and supported by the Clinical and Translational Science Awards (CTSA) grant UL1TR004419 from the National Center for Advancing Translational Sciences. Research reported in this publication was also supported by the Office of Research Infrastructure of the National Institutes of Health under award number S10OD026880 and S10OD030463. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

Author contributions

J.M.D., G.N.N., and A.S. conceptualized the study and curated the data. J.M.D. performed formal analysis. J.M.D., Z.W.H., P.A., G.N.N., and A.S. developed the methodology. J.M.D., G.N.N., and A.S. developed software. G.N.N. and A.S. provided resources and supervised the study. A.S. acquired funding. J.M.D., Z.W.H., and A.S. prepared the original manuscript draft. J.M.D., Z.W.H., M.S. A.S.S., J.G., A.C.C., G.K., R.S., A.G., P.M., V.N., D.P., A.C., D.G., S.A., U.G., M.A.L., R.V., F.F., R.F., A.S.K., A.W.C., I.H., L.C., D.R., P.K., R.K., M.K., P.A., J.A.K., G.N.N., A.S. contributed to the critical revision of the manuscript for important intellectual content. All authors, J.M.D., Z.W.H., M.S. A.S.S., J.G., A.C.C., G.K., R.S., A.G., P.M., V.N., D.P., A.C., D.G., S.A., U.G., M.A.L., R.V., F.F., R.F., A.S.K., A.W.C., I.H., L.C., D.R., P.K., R.K., M.K., P.A., J.A.K., G.N.N., A.S., have read and approved the manuscript.

Competing interests

GLUCOSE is the subject of a provisional patent application (Application No. 63/698,447) filed with the United States Patents and Trademarks Office, in which JMD, AS, and GNN are named inventors. G.N.N. is a founder of Renalytix, Pensieve, Verici and provides consultancy services to

AstraZeneca, Reata, Renalytix, Siemens Healthineer and Variant Bio, serves a scientific advisory board member for Renalytix and Pensieve. He also has equity in Renalytix, Pensieve and Verici. GNN is also an Associate Editor for npj Digital Medicine. He had no role in editorial decisions about the manuscript. L.C. is a consultant for Vifor Pharma INC and has received honorarium from Fresenius Medical Care. J.A.K. reports receiving consulting fees from Astute Medical/bioMerieux, Astellas, Alexion, Chugai Pharma, Novartis, Mitsubishi Tenabe and GE Healthcare and is a full time employee of Spectral Medical. All remaining authors have declared no conflicts of interest.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41746-025-01709-9>.

Correspondence and requests for materials should be addressed to Ankit Sakhuja.

Reprints and permissions information is available at <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) 2025