# An artificial intelligence decision support system for the management of type 1 diabetes

**Nichole S. Tyler, BS**[1], **Clara M. Mosquera-Lopez, PhD**[1], **Leah M. Wilson, MD**[2], **Robert H. Dodier, PhD**[1], **Deborah L. Branigan, BS**[2], **Virginia B. Gabo, BS**[2], **Florian H. Guillot, BS**[2], **Wade W. Hilts, MS**[1], **Joseph El Youssef, MD**[2], **Jessica R. Castle, MD**[2], **Peter G. Jacobs, PhD**[1]

[1]Artificial Intelligence for Medical Systems (AIMS) Lab, Oregon Health & Science University, Portland OR, USA

[2]Harold Schnitzer Diabetes Health Center, Oregon Health & Science University, Portland OR, USA

## Abstract

Type 1 diabetes (T1D) is characterized by pancreatic beta cell dysfunction and insulin depletion. Over 40% of people with T1D manage their glucose through multiple injections of long-acting basal and short-acting bolus insulin, so called multiple daily injections (MDI).[1,2] Errors in dosing can lead to life-threatening hypoglycaemic events ($< 70$ mg dL^−1) and hyperglycaemia ($> 180$ mg dL^−1), increasing the risk of retinopathy, neuropathy, and nephropathy. Machine learning (*artificial intelligence*) approaches are being harnessed to incorporate decision support into many medical specialties. Here we report an algorithm that provides weekly insulin dosage recommendations to adults with T1D using MDI therapy. We employ a unique virtual platform[3] to generate over 50,000 glucose observations to train a K-nearest-neighbours[4] decision support system (KNN-DSS) to identify causes of hyperglycaemia or hypoglycaemia and determine necessary insulin adjustments from a set of 12 potential recommendations. The KNN-DSS algorithm achieves an overall agreement with board-certified endocrinologists of 67.9% when validated on real-world human data, and delivers safe recommendations per endocrinologist review. A comparison of physician-recommended adjustments to insulin pump therapy indicates full agreement of 41.2% among endocrinologists, which is consistent with previous measures of inter-physician agreement (41–45%)[5]. *In silico*[3,6] benchmarking using a platform accepted by the

U.S. Food and Drug Administration for evaluation of artificial pancreas technologies, indicates substantial improvement in glycaemic outcomes after 12-weeks of KNN-DSS use. Our data indicate that the KNN-DSS allows for early identification of dangerous insulin regimens and may be used to improve glycaemic outcomes and prevent life-threatening complications in people with T1D.

---

Optimal management of type 1 diabetes requires precise insulin administration to maintain glucose within safe ranges. Dosage regimens are complicated by day-to-day changes in insulin sensitivity, which can cause large excursions in glucose. Failure to dose insulin properly can result in diabetic ketoacidosis and hypoglycaemia, which may lead to coma or death. Intensive insulin regimens can enhance glycaemic outcomes in people with T1D who use MDI therapy[7], but a number of factors confound adherence to insulin dosing. Fear of hypoglycaemia, challenges with numeracy to calculate meal or correction boluses, changes in insulin sensitivity during exercise, illness, stress andmenstruation, and the psychological toll of this chronic disease make it difficult for people with T1D to adhere to these regimens.[8–11]

Whereas many smartphone apps are available to help people better manage their diabetes, most of these are not validated and have not shown clinical efficacy. A recent review indicated that out of hundreds of such applications, only 12 were validated in clinical trials and few of these significantly improved glycated haemoglobin (HbA1c) in people with T1D.[12–15] Apps shown to improve glycaemic outcomes provided users with weekly or biweekly feedback from health professionals on insulin dosage adjustments.[12,13] Continuous glucose monitoring (CGM) has been shown to significantly improve HbA1C, but as a sole intervention does not bring everyone to goal.[16] CGM-informed advisory systems[17] range from machine learning to physiologic models and heuristic approaches for the adjustment of basal[18–20] and bolus therapies.[21,22] Nimri et al. developed a system to guide adjustment of insulin pump settings using capillary blood glucose or CGM data.[5] Perez-Gandia et al. developed a predictive neural-network to assist with real-time insulin administration or carbohydrate consumption.[23] MDI-inclusive approaches include a model-based decision support system for titration of insulin prior to exercise by Breton et al.[24], and an adaptive KNN case-based reasoning approach for titration of short-acting insulin by Reddy et al.[25]

The KNN-DSS that we describe provides up to four optimally selected dosing and behavioral recommendations once per week to adults with T1D who use MDI therapy. Recommendations are selected to manage insulin dosed for meals andsnacks to bring glucose to within a target range. A virtual patient simulator platform was implemented to design the KNN-DSS algorithm. The virtual patient simulator[3] is a mathematical representation of the glucoregulatory response to food, insulin, and exercise in people with T1D. The virtual patient simulator was used to train a machine-learning KNN[4] model to predict optimal insulin recommendations that improve glycaemic outcomes. Input data for the algorithm are acquired from CGM data, insulin data obtained from Bluetooth-enabled capture devices, and physical activity metrics obtained through wearable sensors (Figure 1a). The KNN-DSS then classifies glycaemic features and delivers recommendations to improve percent time in ange ($70–180$ mg dL$^{-1}$) and reduce percent time in hypoglycaemia (Figure

1b). User-specific titration of insulin occurs using an adaptive learning postprandial hypoglycaemia avoidance (ALPHA) algorithm[26] which selects the optimal bolus insulin based on the prior glycaemic outcomes of the user (Figure 1c). To ensure recommendations conform to physician standards, we developed an expert-knowledge quality control algorithm (Figure 1d, Extended Data 1–7). The heuristic Quality Control algorithm is designed for user safety and operates independent of the machine-learning framework. The KNN-DSS system delivers one or more recommendations from a set of 12 unique recommendations for insulin adjustments and dosage behaviors with respect to long-acting basal insulin, fast-acting carbohydrate-to-insulin ratio (carb:ins), and correction insulin dosage (Table 1). The top three meal and basal insulin recommendations are selected by KNN classification, whereas recommendations for correction doses and compliance with care are supplied by the heuristic ALPHA and quality control algorithms.

We validated the accuracy and safety of KNN-DSS generated recommendations compared with those of board-certified endocrinologists using 687 days of real-world data collected from 25 adult participants on MDI therapy. We demonstrated efficacy of the KNN-DSS during two 52-week studies *in silico* and characterized the response of the engine to dynamic disturbances in glycaemic patterns. Lastly, we report the results of a short, proof-of-concept, single-centre clinical study to evaluate the safety of the KNN-DSS in human participants (Supplementary Table 1: Study 1).

To compare the recommendations of the KNN-DSS with endocrinologist recommendations, we used data collected from 25 adult participants with T1D during a 28-day outpatient study (Supplementary Table 1: Study 1). One of three endocrinologists from Oregon Health and Science University (OHSU) medical center analysed the glucose and insulin dosing data from each participant and provided recommended adjustments to basal insulin, and fast-acting meal insulin and correction insulin during four different windows of time (7:00 – 11:00, 11:00 – 15:00, 15:00 – 20:00, or 20:00 – 7:00) (Table 1). Recommendations regarding daily bolus calculator use were also provided. The KNN-DSS recommendations were labelled as being in full agreement, partial agreement, full disagreement, or partial disagreement with the physician. The accuracy of recommendations delivered by the KNN-DSS as compared to those of board-certified endocrinologists was quantified using a modified Sorenson-Dice coefficient[27] (equation (6)). We measured a combined agreement of 67.9% between endocrinologist and KNN-DSS recommendations, whereas 6.4% of recommendations were in disagreement. In 16.7% of recommendations the engine identified an issue and physician did not indicate a recommendation, and 9% of recommendations were not comparable (Table 2). We performed additional analysis on a subset of these participants who exhibited consistent use of the bolus calculator and adherence to their insulin dosage settings (greater than 75%). For this subset of participants, the engine recommendations were in full agreement with physicians 50.8% of the time and exhibited an overall agreement of 67.5% with the recommendations of physicians. We observed that over 99% of recommendations delivered across the 100 weeks of data by the KNN-DSS passed a safety review in which the endocrinologists reviewed each recommendation for the potential to cause hypoglycaemia episodes and overnight events.

The measure of physician agreement with the KNN-DSS is similar to those found in published studies involving other decision support systems. In an international, multicenter study,[5] Nimri et al. compared physician-recommended adjustments of insulin pump settings and found the expected full agreement of different physicians to be between 41% and 45%, and the expected full disagreement between 9% and 12%. The results demonstrated high variability among physicians both internationally and in practice at the same institution. Nimri et al. used multiple study centers and evaluated insulin pump settings, which differentiated their study from that of a single-center study on people using MDI. Additional studies that compared adjustments in MDI therapy reported that general practitioners and endocrinologists alike identified 67% of indicated changes to insulin,[28] and reported an agreement of 63% between physician-recommended and software-recommended titrations to insulin dosages.[29]

We measured inter-physician recommendation variability on a dataset collected previously from participants with T1D using sensor augmented pump therapy[30] (Supplementary Table 1: Study 2). We found a value of 41.2% in relation to full agreement among endocrinology faculty at OHSU (Supplementary Table 2). The KNN-DSS engine, trained with a virtual platform, demonstrates high full agreement (50.8%) and partial agreement (67.5%) with endocrinologist recommendations when validated on real-world data (Study 1), and exceeds inter-physician agreement found in Study 2.

We evaluated the ability of the KNN-DSS to improve glycaemic outcomes using two virtual patient simulators[3,6]. Each virtual patient was given real-world meal scenarios previously recorded during a clinical trial of automated insulin delivery therapy. These meal scenarios are meant to rigorously challenge algorithm performance with realistic eating patterns.

Evaluation *in silico* demonstrated the ability of the KNN-DSS to identify problematic glycaemic patterns and to deliver effective insulin dosage recommendations. In the first *in silico* study using the OHSU T1D simulator[3], 29 virtual patients with varying adherence to insulin therapy, varying circadian insulin sensitivities and carbohydrate misestimation of $\pm 30\%$, were evaluated in a 75-week study of weekly decision support. After 12 weeks of use, the KNN-DSS and supporting algorithms improved virtual patient outcomes considerably, increasing average percent time in target range from 59.5% to 79.8% (P = 2E-5), maintaining percent time in hypoglycaemia at target ($<2\%$) and reducing inter-individual variability (Figure 2a, Supplementary Table 3).

After the optimal insulin dosing settings were obtained at 52 weeks of simulation, we disturbed the system by imposing new insulin dosing errors and changes to patient settings (Supplementary Table 4). The engine was able to correct the problems with the invalid dosing settings, gradually improve the time in target range, and reduce the time in hypoglycaemia (Figure 2b). Patient settings largely trended towards their pre-disturbance values (Figure 2c), which indicates the engine's ability to respond to dynamic changes and retrieve the original settings. Since there are many possible combinations of long-acting and short-acting insulin therapy that may improve glycaemic control in a person with T1D, not all therapy settings returned to the original settings prior to the disturbance. We nonetheless

found that all of the glycaemic outcomes of the virtual patients still improved following the disturbance (Figure 2b).

In the second 52-week *in silico* study, we used a benchmarking platform accepted by the United States Food and Drug Administration for the evaluation of artificial pancreas algorithms, the UVA-Padova simulator[6], to evaluate the performance of the KNN-DSS on 100 virtual adult patients, 100 virtual adolescent patients, and 100 virtual pediatric patients. In adult patients, percent time in target range improved from 75.1% at baseline to 81.8% (P=1E-4) at the study conclusion. Percent time in hypoglycaemia was reduced from 4.0% at baseline to 0.55% (P=2E-12) at the end of the study (Supplementary Table 3, Extended Data 8). *In silico* studies may provide optimistic estimations regarding glycaemic outcomes because virtual study patients exhibit near perfect adherence to recommended insulin adjustments, which often does not happen in real-world studies. Nonetheless, the results shown here indicate that study participants who use the recommendations provided by the KNN-DSS may expect a reduction in hypoglycaemia and an improvement in time in target range after 12 weeks. Notably, at 4 weeks, both the UVA-Padova and the OHSU T1D simulators showed reductions in glycaemic variability and hypoglycaemia and a small average improvement in glycaemic time in target range, with many patients showing a reduction in time in target range. The KNN-DSS is designed to prioritize safety and reduce hypoglycaemia as well as optimize insulin dosages and percent time in target range; as a consequence, time in target range may initially be reduced as problematic and aggressive insulin doses are titrated down, and improves significantly after 12 weeks of continued use. Performance was also evaluated using the UVA-Padova simulator in an adolescent and pediatric population. For adolescents and children, we also observed significant improvements in glycaemic outcomes, as percent time in target range increased from 68.2% to 75.2% (P = 8E-9) for adolescents, and 65.7% to 69.1% (P = 0.02) for children after 12 weeks (Supplementary Table 5).

We evaluated the safety of the KNN-DSS in a single-center feasibility study. Sixteen adults out of 25 adults with type 1 diabetes on MDI therapy (Supplementary Table 1: Study 1) underwent weekly KNN-DSS augmented decision support. These 16 participants were given weekly recommendations on dosing and behavioral changes from a physician who reviewed the glucose history of participants and recommendations suggested by the KNN-DSS. Although this small pilot study was not powered to detect any impact from the intervention, and the study duration of 4 weeks was too short to observe a significant impact on time in target range, we did observe that participants exhibited a 25% decrease in hypoglycaemia events over a 24-hr period (from 0.86 to 0.64 events per day, P = 0.051), a 33% decrease in daytime hypoglycaemia events (from 0.43 to 0.29 events per day, P = 0.096), a 43% decrease in hypoglycaemia overnight (from 0.50 to 0.29 events per day, P = 0.04) and an 76% decrease in serious hypoglycaemia (<54 mg dL^−1) overnight (from 0.48% to 0.11%, P=0.03) during the final week compared with the first week in the study (Extended Data 9). Fifteen of the 16 participants completed all 4 weeks of the study whereas one participant chose to exit the study after two weeks. In a similar manner, we observed no substantial changes in daytime percent time in target range comparing the final week to week 1 (54.2% to 52.6%, P = 0.63), a decrease in time in target range overnight (56.7% to 45.7%, P = 0.034), a non-significant decrease in time in target range over a 24-hr period (55.3% to

49.1%, P = 0.06), and a non-significant increase in mean glucose (172.4 mg dL^−1 to 185.5 mg dL^−1, P = 0.08). As shown in Supplementary Table 3, our simulation studies confirm the findings of the clinical study, which indicates that, although a reduction in hypoglycaemia can be expected after four weeks, a longer study is required before we can expect substantial improvements in time in target range.

Other groups have also reported on the impact of automated decision support systems on glycaemic outcomes, with results primarily indicating that hypoglycaemia can be reduced. Breton et al. recruited 24 participants with T1D (8 of whom were on MDI therapy) for a crossover study to evaluate automated decision support. Glycaemic outcomes were evaluated during a 48-hour in-patient session in which participants underwent standardized meals and exercise. They observed a statistically significant reduction of percent time in hypoglycaemia for the DSS vs. the control group (3.2% [1.3, 4.8] vs. 0.9% [0.4 2.3], P = 0.02) but no significant change in percent time in target range.[24] Herrero et al. described a case-based reasoning DSS (named ABC4D) and evaluated the system *in silico* using 20 representative adults and adolescents from the UVA-Padova simulator population. They found that after 4 weeks, percent time in target range could be improved in adults from 75.2 ± 11.7 to 81.9 ±13.4 (P<0.05) and percent time in hypoglycaemia could also be reduced from 0.3±0.5 to 0 (P=0.17).[22] Although we also used the UVA-Padova population, a direct comparison of ABC4D to the KNN-DSS would be difficult as we evaluated our algorithm in 300 UVA patients. Reddy et al. further evaluated the ABC4D algorithm in a real-world, 6-week pilot study, but whereas both percent time in target range and percent time in hypoglycaemia improved slightly, the changes were not significant.[25] Other approaches described a reduction in hypoglycaemia after 12 weeks of basal and fast-acting insulin decision support, but no significance measures were reported[5]. A common theme of these studies is that it has not yet been shown that a DSS can improve percent time in target range in human studies. Our *in silico* results indicate that longer study durations (>12 weeks) will be necessary to demonstrate improvements in percent time in target range through use of the KNN-DSS.

The KNN-DSS was trained and validated for use with specific sensor technologies, insulin therapies, and target populations. The KNN-DSS system utilizes CGM devices that sample glucose at 5-min intervals. Flash glucose systems are compatible with the KNN-DSS algorithm; however these will require further testing to handle the asynchronous sample rates of flash glucose systems. Like flash glucose monitoring systems, fingerstick glucose is also measured at varying intervals (4–10 times per day), and more training and testing is needed before this could be incorporated into the KNN-DSS. The engine is also compatible with the large majority (~95%) of existing insulin therapies including fast-acting aspart or lispro and long-acting Lantus (glargine), Tresiba (degludec), and Toujeo (glargine U300) basal formulations. Intermediate-acting NPH insulin (which represents <5% of use cases) will require additional testing and evaluation. Although we report on adults, adolescents and children in this article, further work will be needed to assess performance in vulnerable and complex populations, including the elderly and pregnant women. Moreover, virtual simulators have not been developed to fully represent these populations, which makes it challenging to incorporate these populations into the design. The KNN-DSS will need further training before it can be targeted to these groups.

We explored how specific glucose and insulin features were related to optimal recommendations calculated by the KNN-DSS algorithm (Supplementary Table 6). We observed that the KNN-DSS mapping of specific glycaemic features to optimal recommendations matches intuition and, in general, is synonymous with physician opinions regarding titration of insulin dosing for people with T1D.

We have shown that an artificial intelligence decision support system with an expert-knowledge quality control algorithm can be used to help people with T1D identify problematic glycaemic patterns at the same level of accuracy as board-certified endocrinologists. Our unique *in silico* training platform enables us to generate training sets of diverse glycaemic profiles from the OHSU T1D simulator. The final engine design performs well on independent *in silico* virtual populations and, most notably, can identify insulin dosage issues in real-world human data. Further validation in longer clinical trials in humans is critical to understand how artificial intelligence-based decision support systems can improve glycaemic outcomes in people with T1D.

## Methods

### K-nearest-neighbours design

The K-nearest-neighbor classification algorithm (KNN)[4] is a supervised machine learning approach that matches input features with an outcome variable or class. We define the KNN input features as specific glycaemic outcomes, such as percent time in target range (70–180 mg dL^−1), percent time spent in hypoglycaemia (<70 mg dL^−1), the number of meal- or correction-related hypoglycaemia episodes and so on (see Supplementary Table 7 for complete list of features). The outcome variable or class that the KNN predicts is a recommended adjustment to insulin dosage that leads to an improved percent time in target range and a reduction in percent time in hypoglycaemia. This classification is accomplished using a look-up table that matches the weekly glycaemic features of a person with optimal dosing recommendations. The KNN approach identifies unique recommendations regarding long-acting basal insulin and fast-acting meal insulin. Recommendations regarding adherence and correction factors are accomplished separately by the ALPHA and Quality Control algorithms described in detail below (Figure 1).

### Training Dataset generation

We generated the look-up table of optimal recommendations using an *in silico* virtual patient simulator consisting of 99 individuals with T1D exhibiting diverse glycaemic dynamics generated by variations in insulin sensitivity and daily insulin requirements, carbohydrate sensitivity, weight differences, circadian insulin sensitivity and variations in adherence to insulin dosing[3]. Circadian insulin sensitivity was achieved through continuous modulation of insulin-mediated peripheral glucose uptake, peripheral insulin uptake, and hepatic glucose production, varying within 20% of their original values[3]. CGM data and insulin data were obtained from 70 of the virtual patients over the course of a 15-week *in silico* study, while the other 29 patients were retained for an evaluation of algorithm performance in a separate *in silico* study. Real-world meal scenarios provided to the virtual test subjects were obtained from a previous clinical trial[31] in which we acquired 80 d of realistic meal patterns and

carbohydrate content. Forty of these daily meal scenarios were used to train the KNN-DSS, while the other 40 were used to validate the algorithm. These daily meal-scenarios were randomized and administered to virtual patients and the Pettus-Edelman[32] approach for CGM trend arrow adjustment and a bolus calculator was used to dose mealtime insulin. The optimal recommendations for an individual were identified by simulating the glycaemic outcomes from administering each dosing recommendation given in Table 1. The glycaemic outcomes of time in glucose target range (70–180 mg dL^−1), mean glucose, hypoglycaemia (<70 mg dL^−1), and serious hypoglycaemia (<54 mg dL^−1 ) were measured the week following each simulated dosing scenario. During the training of the KNN, these measured outcomes were used to define a heuristic objective function to select the optimal recommendations, {$R$_Opt}, to maximize the measured percent time in target range, and reduce percent time in hypoglycaemia ($X$_TIR, $X$_Hypo in equation (1), respectively). We provide this heuristic objective function in equation (1). In step 1, we identify a subset of recommendations, {$R$_Hypo}, that yields percent time in hypoglycaemia less than or equal to 2% and no serious hypoglycaemia. If the patients exhibit persistent hypoglycaemia, we identify the recommendations that minimize hypoglycaemia, {$R$_Hypo}. Recommendations that yield a percent time in hypoglycaemia greater than 2%, or any extreme hypoglycaemia, are excluded from the subset of optimal recommendations. In step 2, we identify from this subset the optimal recommendation, { $R$_Opt}, which yields the largest improvement in the time spent in target glucose range, or mean glucose if the percent time in target range is too small due to persistent hyperglycaemia.

$$R_{Opt} =$$

$$
\begin{cases}
\textbf{Step 1a}, \ \{R_{Hypo}\} = \bigcup_{i=1}^{n} i * \delta(x_{hypo}) * \gamma(x_{serious\ hypo}) & \quad \delta(a) = \begin{cases} 0, a > 2 \\ 1, a \leq 2 \end{cases} \\[2mm]
 & \quad \gamma(a) = \begin{cases} 0, a \neq 0 \\ 1, a = 0 \end{cases} \\[2mm]
Step\ 1b, \ \{R_{Hypo}\} = \mathrm{argmin}_{X_{Hypo}} \sum_{i=1}^{n} x_i & \rightarrow perform\ if\ step\ 1a\ yeilds\ empty\ set \\[2mm]
\textbf{Step 2a}, \ \{R_{Opt}\} = \mathrm{argmax}_{X_{TIR}} \sum_{i=R_{Hypo}} x_i \\[2mm]
Step\ 2b, \ \{R_{Opt}\} = \mathrm{argmin}_{X_{MeanGlucose}} \sum_{i=R_{Hypo}} x_i & \rightarrow \ perform\ if\ step\ 2a\ yeilds\ empty\ set
\end{cases}
$$

(1)

This optimal recommendation was stored with the glycaemic features from the prior week to form an observation in the look-up table. The observation of paired weekly glycaemic features and optimal insulin dosage were compiled into a look-up table used in the KNN algorithm. Additional real-world behavioral scenarios were imposed during the *in silico* study by programming the virtual patient to perform one or more of 13 insulin dosing errors (Supplementary Table 4), to periodically accept or fail to accept the advice of the recommendation given each week and to use one of two different types of bolus calculators. The size of the final look-up table totaled 51,831 observations.

### K-nearest neighbours parameter identification and feature selection

The features used in the KNN-DSS included features drawn from CGM, physical activity data, and long-acting and short-acting insulin data. We identified an optimal set of features through a feature selection technique called *'greedy'* sequential forward selection33. We determined the optimal number of neighbours by performing a grid search using the optimal set of features. The final KNN design used 30 neighbours and 25 glycaemic features to perform classification (see Supplementary Table 7). The weighting scheme was decided by comparing classification accuracy across (1) no weighting, (2) distance-based weighting, (3) class-based weighting, and (4) combined distance and class-based weighting (Supplementary Tables 8–9).

### Feature Importance

We evaluated which features contributed most significantly to each recommendation by calculating the '*mutual information*' between each feature and the recommendation selected by the classifier. The mutual information between two random variables (for example, a feature *f* with a distribution F and a recommendation *r* with a distribution R) with a joint probability mass function P_FR) is defined according to equation (2).

$$I(F, R) = \sum\nolimits_{f, r} P_{FR}(f, r) \log\left(\frac{P_{FR}(f, r)}{P_F(f)P_R(r)}\right) \tag{2}$$

The mutual information of features was calculated for each separate recommendation in the classifier. A one-vs-all approach was used, in which the mutual information considers a single recommendation class as positive and all other classes as negative. This was repeated for each recommendation class in order to generalize what features contribute most significantly to a given recommendation class. Relative feature importance determined for each recommendation is listed in Supplementary Table 6.

### Precise Insulin Titration

The KNN-DSS assumes that meal insulin doses are calculated using carb:ins ratio, and correction boluses are calculated correction factors and glucose trends[32,34]. The also uses smart-bolus calculations that incorporate meal-time corrections and active insulin on board (IOB) (equation (3)).

$$\begin{cases} Recommended\ Insulin\ Bolus = Bolus_{meal} + Bolus_{correction} - IOB \\ Bolus_{meal} = Grams\ of\ Carbohydrate\ x\ \dfrac{1\ unit\ of\ insulin}{carb:ins} \\ Bolus_{correction} = (Current\ Glucose + Offset - Target\ Glucose)\ x\ \dfrac{1\ unit\ of\ insulin}{correction\ factor} \end{cases} \tag{3}$$

Precise titration of insulin is accomplished using heuristic approaches that adjust carb:ins ratios and correction factors. The ALPHA algorithm, described in our recent publication[26], retrospectively analyses the average glycaemic response of a person to insulin boluses and returns an aggressiveness factor (*A_f*). The framework of this algorithm, as well as the

modifications to the original algorithm for the current paper implementation, is described as follows. Separate analysis is performed for both meal-related boluses and correction boluses, again across different windows of time (see Table 1). For each bolus entry, ALPHA adapts the aggressiveness factor if the glucose of the person is outside of target range following meals or corrections. The aggressiveness factors corresponding to each bolus are then used to calculate an average, smoothed aggressiveness factor ($A\_f\hat{A}vg$) shown in equation (4). This smoothed aggressiveness factor is used to adjust carb:ins ratio and correction factor settings.

$$A_f^{Avg}(k) = \frac{A_f(k) + A_f^{Avg}(k-1) + A_f^{Avg}(k-2)}{3} \tag{4}$$

In the implementation discussed herein, the aggressiveness factor assigned to an individual bolus delivered at time k, $A\_f(k)$ is determined using a piece-wise linear adjustment that is a function of the minimum glucose ($G\_min$) measured within 4 h of the last meal bolus or within 3 h after the last correction bolus (equation (5)).

$$A_f(k) = \tag{5}$$

$$\begin{cases} 0.4, & 0 \le G_{min} \le G_{hypo} \\ \frac{(G_{min} - G_{hypo}) * (A_f^{Avg}(k-1) - 0.4)}{G_{eug-lower} - G_{hypo}} + 0.4, & G_{hypo} \le G_{min} \le G_{eug-lower} \\ A_f^{Avg}(k-1), & G_{eug-lower} \le G_{min} \le G_{eug-upper} \\ \frac{(G_{min} - G_{eug-upper}) * (1.3 - A_f^{Avg}(k-1))}{G_{hyper} - G_{eug-upper}} + A_f^{Avg}, & G_{eug-upper} \le G_{min} \le G_{hyper} \\ 1.3, & G_{hyper} \le G_{min} \end{cases}$$

If $G\_min$ is within the target range of $G\_eug-lower$ = 90 to $G\_eug-upper$ = 140 mg dL^−1 (where $G\_eug-lower$ and $G\_eug-upper$ are the lower and upper limits of a euglycaemic glucose target range, respectively), then the aggressiveness factor does not change and $A\_f$ = $A\_f\hat{A}vg$. However, if $G\_min$ drops below $G\_eug-lower$, the aggressiveness factor, $A\_f$, is reduced proportionally down to a hypoglycaemia threshold of $G\_hypo$ (70 mg dL^−1 ). Below the hypoglycaemia threshold ($G\_hypo$), $A\_f$ = 0.4 which means that the pre-meal insulin will be dosed at 40% of the original amount as shown in equation (5). In a similar manner, the aggressiveness factor is increased proportionally with respect to $G\_min$ if $G\_min$ is above the upper limit of target range ($G\_eug-upper$) until $G\_min$ exceeds the hyperglycaemic threshold ($G\_hyper$). Above $G\_hyper$, the value of the insulin aggressiveness factor is 1.3. To ensure that the aggressiveness factor accurately reflects user

glycaemic response, a minimum of 5 boluses must be observed within a specific window of time before a new aggressiveness factor is calculated.

Adjustment of carb:ins and correction factor occurs under two separate scenarios. The ALPHA algorithm is used to calculate the precise dosage adjustment to carb:ins only when indicated by the KNN classification procedure. In contrast, the ALPHA algorithm adjusts correction factors directly because the KNN classification does not address correction factors. The maximum adjustment to fast-acting insulin is constrained to ±15% per week. For adjustments to long-acting basal insulin, the dosage is adjusted by ±10% per week when indicated by the KNN classification procedure. The KNN-DSS system accounts for the extended (> 5 d) pharmacologic steady state of ultra-long-acting Tresiba (degludec) and Toujeo (glargine U300) by constraining basal recommendations to one basal insulin recommendation every 2 weeks (see Extended Data 2–3). These constraints are a safety measure. Titrations to basal insulin are applied uniformly to all basal doses that may occur at different times of day. For example, for people who require twice-daily basal insulin injections, if a 10% reduction in basal insulin is recommended, both the morning and evening insulin will be reduced by 10%.

### Quality Control Algorithm

A quality control algorithm was developed to ensure that recommendations delivered to the person adhere to physician standards. This algorithm incorporates expert knowledge based on physician input on titration of basal and bolus insulin regimens to ensure that engine recommendations are consistent with physician standards and are safe for a person with T1D. Quality control metrics for each recommendation delivered by the KNN-DSS are shown in Figure 1d and elaborated further in Extended Data 1–7.

### Clinical Study Data and Physician Review

Data were obtained from 25 people with T1D who participated in a 4-week, outpatient study of CGM-augmented MDI therapy (Supplementary Table 1: Study 1). After the data collection techniques were optimized on the first 9 participants, the remaining 16 participants received recommendations for dosing and behavioral changes on the basis of suggestions from the KNN-DSS system. At the end of each study week, an endocrinologist reviewed the data and identified one or more adjustments to insulin therapy. Study participants were equipped with a Dexcom G6 sensor and Apple Watch to track glucose trends and physical activity. Participants used either long-acting Lantus or Tresiba insulin that was captured automatically using the Bluetooth enabled Gocap or Clipsulin insulin dose-capture devices. Participants used fast-acting Novolog (aspart) insulin captured during the study automatically with an InPen device. Participants were instructed to log meals and exercise using a custom food and exercise tracking app, and to dose insulin using the InPen app. Out of the 25 participants, 15 were female, the mean weight was 82.73 +/– 19.56 kg, and mean height 170.60 +/– 19.56 cm. The mean duration of diabetes was 15.52 +/– 6.92 years, mean age was 30.50 +/– 5.92 years, and mean A1c was 8.78 +/– 1.36 %. Additional information regarding study population characteristics, recruitment, and ethics oversight can be found in the Reporting Summary.

The study concluded with a total of 78 physician review sessions accounting for over 500 d of data from 25 participants. For each week of study data, physicians were instructed to identify one or more adjustments to insulin therapy from a set of 12 potential recommendations (Table 1). Data obtained during this clinical study were retrospectively analysed by the KNN-DSS to generate recommendations and calculate physician agreement metrics according to Supplementary Table 10.

### Safety Review

Recommendations generated by the KNN-DSS for human participants underwent safety evaluation by faculty at the Department of Endocrinology, Harold Schnitzer Diabetes Health Center. Safety of the KNN-DSS recommendations was assessed by having the physician determine whether the recommendation had the potential to cause hypoglycaemia or night-time hypoglycaemia events. A total of 100 safety reviews were performed.

### Missing Data

At least four days of cumulative CGM data are required by the KNN-DSS framework to provide new recommendations. The KNN-DSS engine framework will refrain from providing a recommendation until sufficient CGM data are present. To address issues of missing data and data misclassification of insulin boluses that are common in real-world datasets, we developed an auxiliary insulin bolus estimation tool. Insulin boluses recorded by the Bluetooth-enabled insulin capture devices were first paired to announced meal entries. Boluses that were not within 20 min of an announced meal were counted as unlabeled boluses. For each unlabeled bolus, we evaluated the glucose level, glucose trend, and correction factor setting at the time of bolus administration. We then estimated the correction insulin dose that would have been called for by inputting this information into the Scheiner trend adjustment calculator[34]. Any remaining units of insulin are counted as a meal bolus. These estimated contributions are then combined with existing insulin data and are input to the algorithm.

### Assessment of the Accuracy of KNN-DSS Recommendations as Compared to Endocrinologist Recommendations

Glycaemic outcomes and insulin dosing behaviors were analysed by one of three board-certified endocrinologists and by the KNN-DSS. Recommendations delivered by the KNN-DSS were compared to physician recommendations for each week of data collected during the clinical study. Similarity between KNN-DSS recommendations and physician recommendations were calculated using a modified Sorensen-Dice similarity coefficient[25].

$$DSC = \frac{\left| R_{Engine} \cap R_{Physician} \right|}{\left| R_{Engine} \right|} \tag{6}$$

In equation (6), the similarity between physician and engine recommendations is calculated as the number of recommendations common to both sets, divided by the total number of recommendations delivered by the engine.

Recommendations were classified into one of three categories: '*agreement*', '*disagreement*', or '*additional treatment*' (Supplementary Table 10). *Agreement* refers to an engine recommendation that was in full agreement with the physician recommendation (a perfect categorical match), or that was in partial agreement with the physician recommendation and titrates insulin in the same direction (for example, different categorical recommendations that both increase insulin). *Disagreement* refers to an engine recommendation that was in full disagreement with the physician (for example, one recommendation increases basal insulin and the other decreases basal insulin), or that partially disagrees with the physician and titrates overall insulin in a different direction (for example, one recommendation increases meal insulin and the other decreases basal insulin). *Additional treatment* refers to a scenario in which the engine recommended insulin dosage adjustments, but the physician indicated no change to the settings of the study participant, and vice-versa. Short-acting insulin bolus treatments reflect a 4 h pharmacokinetic activity; therefore insulin doses in adjacent treatment windows are highly correlated and are considered in partial agreement. A behavioral recommendation to be more adherent to a dosing regimen was counted as safe and in *agreement*. In some scenarios where the engine recommended to use the bolus calculator and the physician recommended to increase or decrease basal insulin, the recommendations are not comparable. The overall accuracy was calculated by averaging the similarity across all recommendations.

### Inter-physician Recommendation Agreement

Using a dataset[30] collected during a 1-month outpatient clinical study of open-loop insulin therapy, three board-certified endocrinologists separately reviewed participant CGM data and dosing behaviors and recommended one or more adjustments to insulin therapy (Supplementary Table 1: Study 2). The physicians then collectively reviewed participant data to reach a consensus on what recommendation should be given. The Sorensen-Dice coefficient (equation (6)) was then used to determine the agreement between individual physician recommendations, as well as the accuracy of physicians' recommendations, as compared to the consensus (Supplementary Table 2). A summary of dataset description and usage is available in Supplementary Table 1.

### *In silico* evaluation

We evaluated the KNN-DSS during two *in silico* studies. In the first study, 29 virtual patients from the OHSU T1D simulator participated in a 75-week study in which the virtual patients used the decision support system weekly to adjust doses to basal insulin, mealtime insulin, and correction insulin. After 52 weeks, we changed certain insulin settings and dosing behaviors and monitored the ability of the engine to recover these settings. In the second study, 100 virtual patients from the UVA-Padova simulator participated in a similar 52-week study of engine usage. To simulate an MDI population using the UVA-Padova simulator, we replaced the default time-varying basal rate, which is characteristic of programmable insulin pumps, with a constant basal dosage that could be titrated weekly by the KNN-DSS.

Virtual patients from both studies exhibited inter-individual variations in weight, total daily insulin requirement, and circadian insulin sensitivities as described above. For both the OHSU T1D and UVA-Padova simulators, patients were fed real-world meal scenarios

(Supplementary Table 1: Study 3) that ranged from 2–9 meals per day and occurred at varying intervals. In addition, we imposed errors in insulin dosing settings and adherence to dosing strategies (see Supplementary Table 4), as well as statistical variation in estimation of meal amounts to reflect realistic use of bolus calculators, interstitial glucose CGM measurement noise and circadian variation in insulin sensitivity to reflect realistic glycaemic profiles. We evaluated study outcomes of percent time in target range and percent hypoglycaemia at time points of 1 week, 1 month, 3 months as well as at the study conclusion.

The virtual patients used for evaluation and a subset of the real-world meal scenarios were excluded from the training process. In this way, performance was analysed on virtual people with T1D that had not been observed before by the KNN-DSS.

## Analysis and Statistical Power

The glycaemic outcomes of percent time in target range and percent time in hypoglycaemia were determined for the virtual patients at each time point of the study. The percent change was calculated across each week of the study compared with the first week of the study, before any recommendations were given. Results are reported by mean and s.d. for normally distributed outcomes, and median and interquartile range for non-parametric data. A students two-tailed, paired t-test of alpha = 0.05 was used to determine the significance in the change of glycaemic outcomes, and a two-tailed Wilcoxon signed-rank test was used to determine significance for non-parametric data. Cohen's d effect size for paired samples was calculated to account for the influence of the *in silico* framework[35] and large sample sizes on p-value statistics.

## Use of Human Subjects

All participants were adults enrolled under informed consent. The pilot study was approved by the Institutional Review Board at OHSU, and additional information can be found at clinicaltrials.gov under registration number NCT03443713.

## Data Availability

The data generated *in silico* during this study and the code used for analysis is available from the corresponding author on request. Access to human participant data was granted for the current study, and further human data usage or sharing is subject to restrictions and is not publicly available. Requests for restricted, de-identified data on human participants can be submitted to the corresponding authors at OHSU. Requests will be assessed on a case-by-case basis, and are subject to a formal Repository Sharing Agreement. Additional reported outcomes of human participants can be found at clinicaltrials.gov under registration number NCT03443713.

## Code Availability

The code used to generate *in silico* data for this study, the OHSU virtual patient population simulator code, is available at https://github.com/petejacobs/T1D_VPP. Access to the licensed software for the UVA-Padova virtual population was granted for the current study

and it can be requested from the developers of this software directly at the University of Padova.
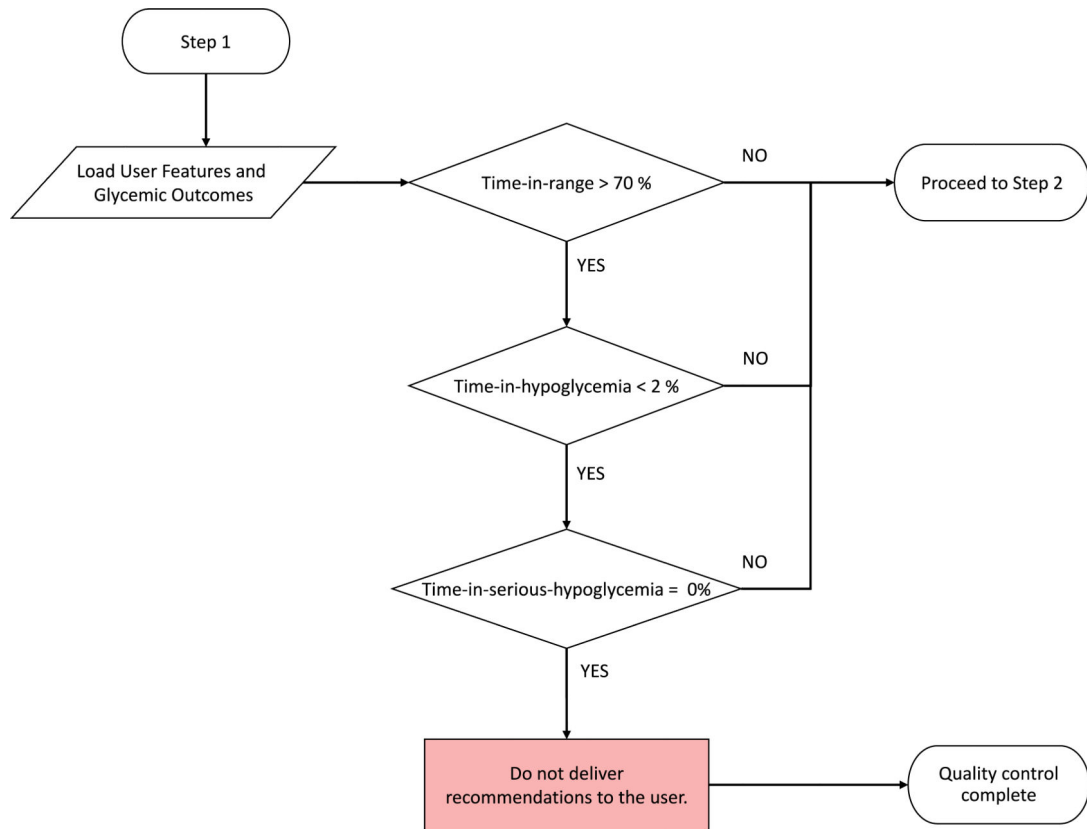
## Extended Data
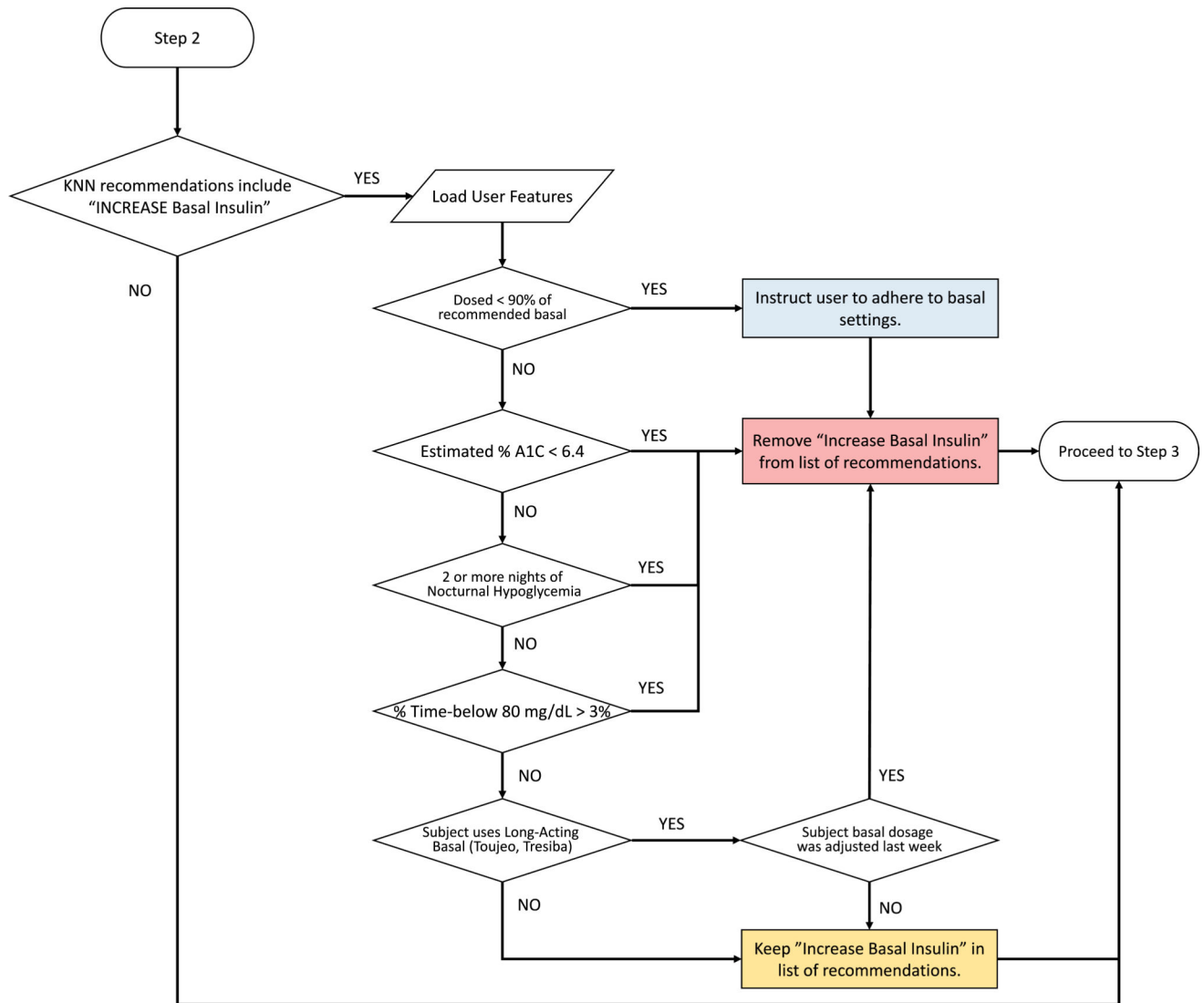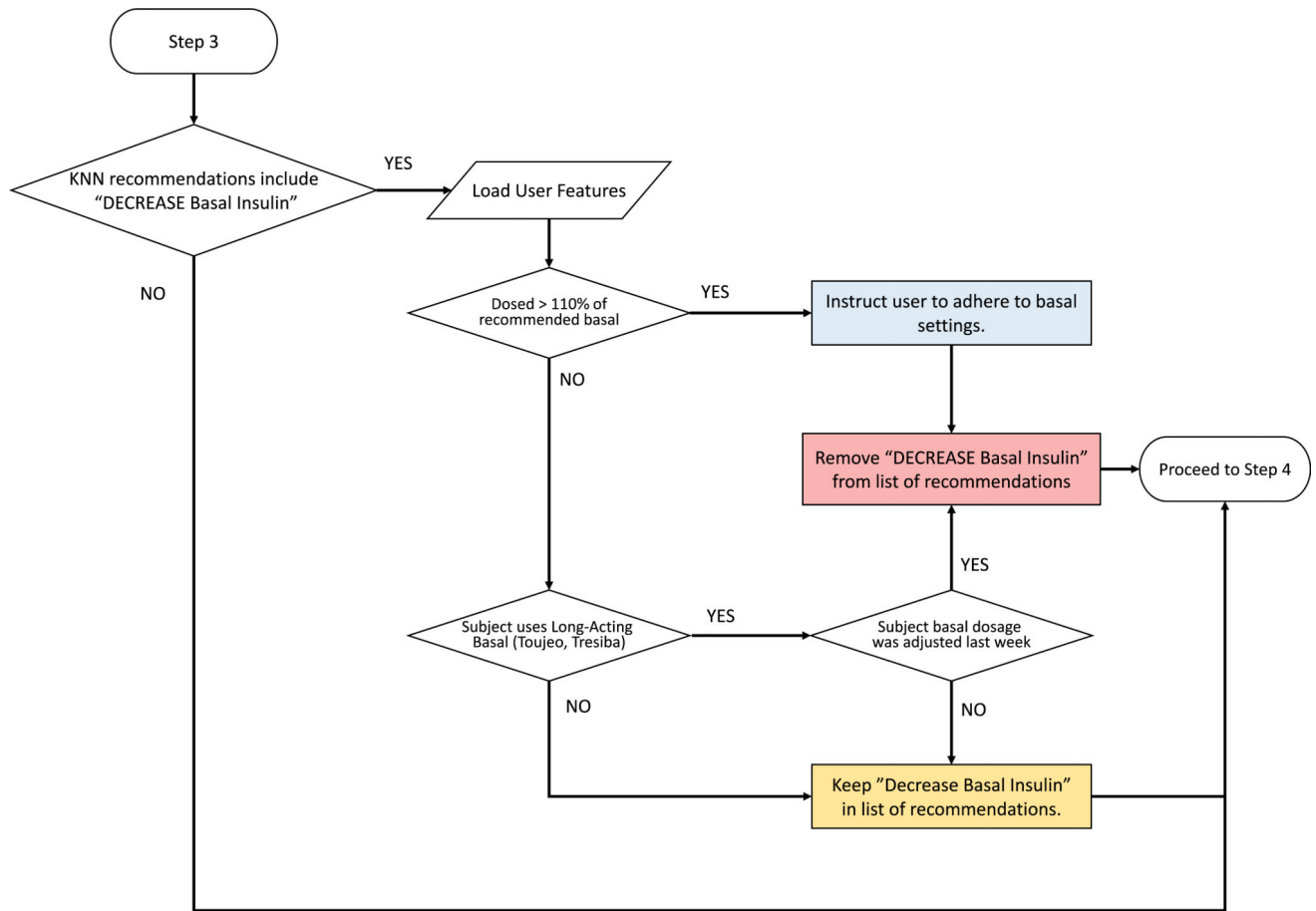
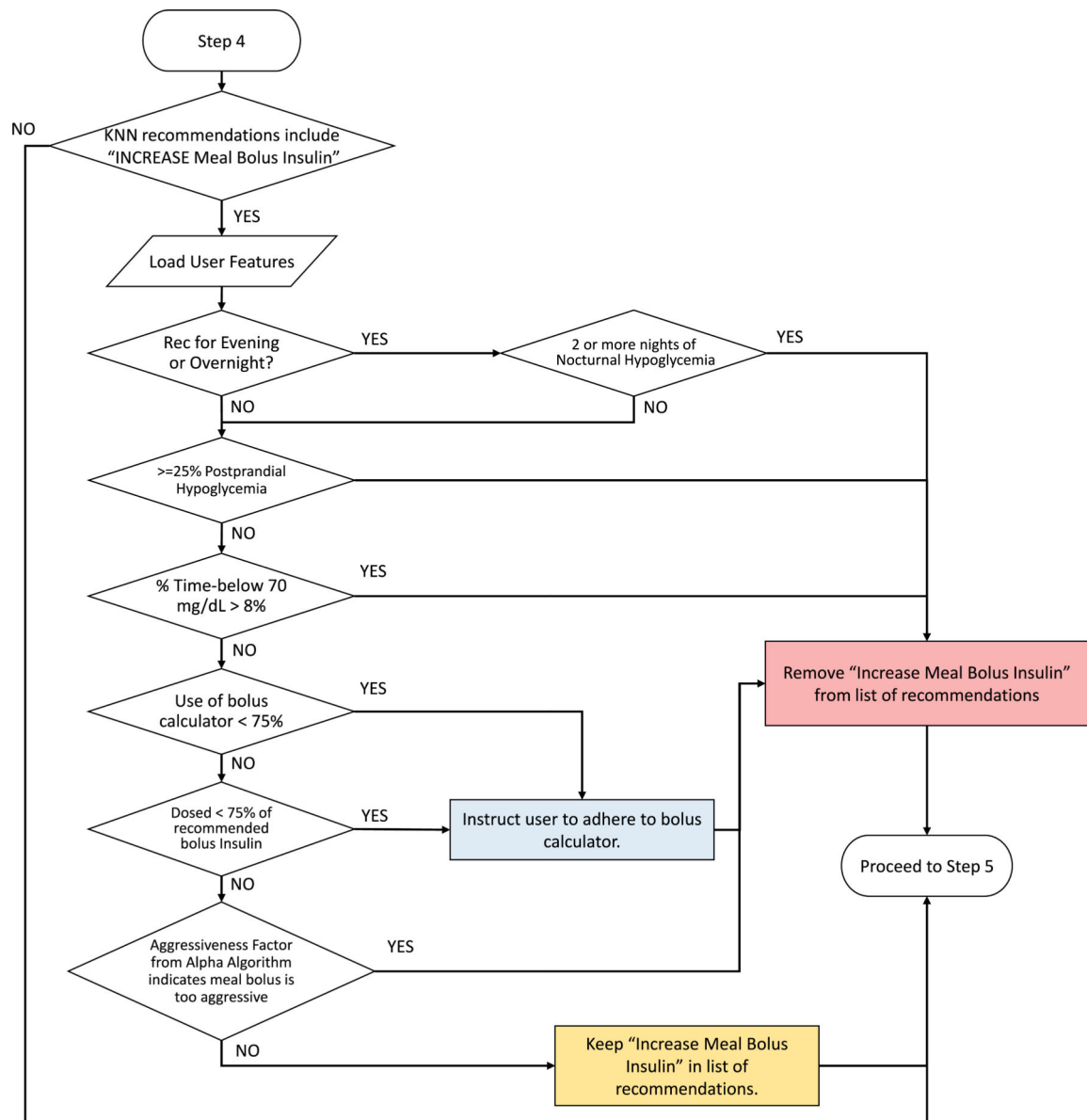**Extended Data Fig. 1. Quality control algorithm to assess need for insulin titration.**
Quality control algorithm to assess need for insulin titration. User data and glycaemic outcomes are loaded and compared against metrics for percent time in hypoglycaemia, percent time in target range, and percent time in serious hypoglycaemia. If users meet all metrics, recommendations for insulin titration are not required.
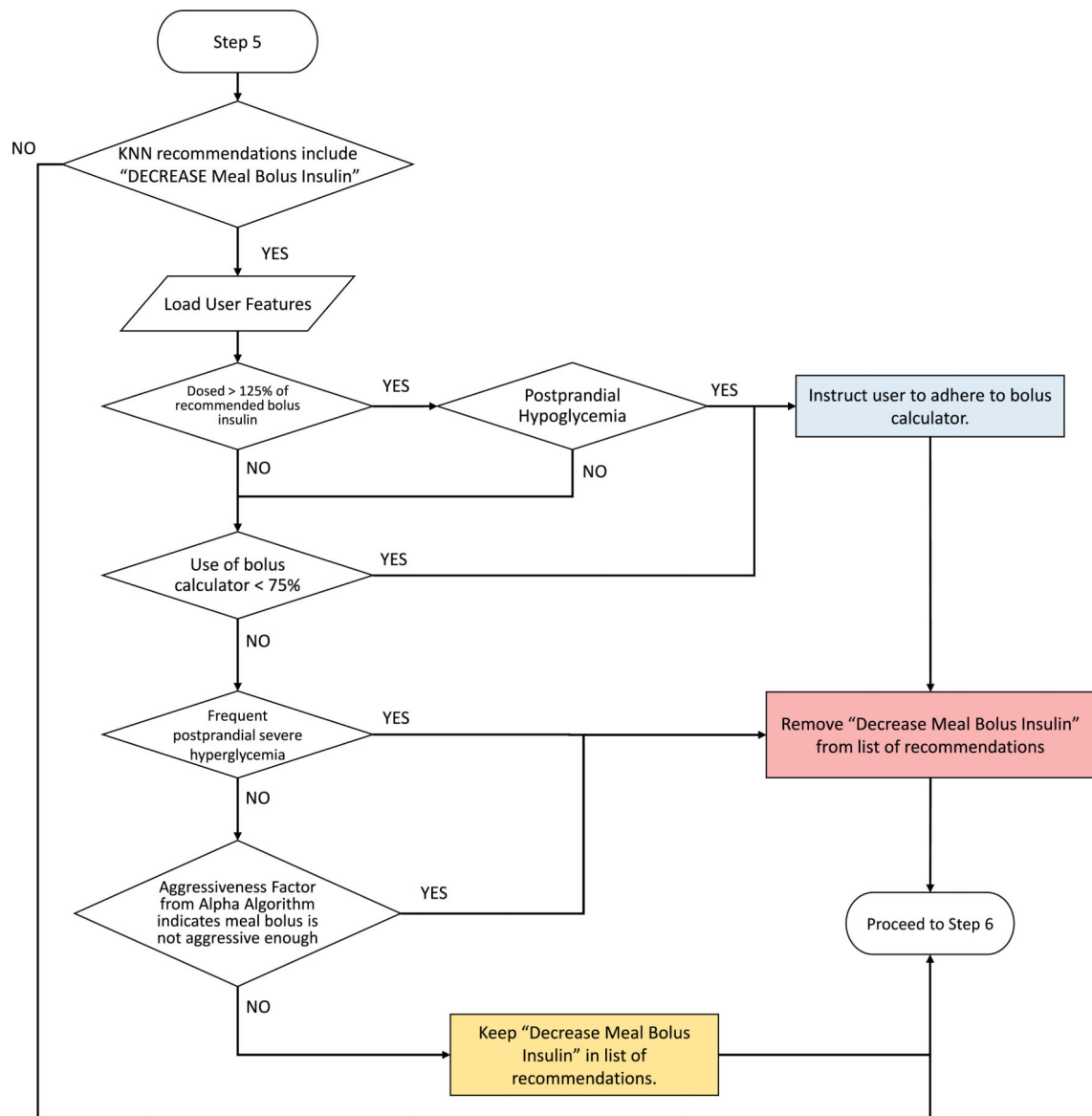
**Extended Data Fig. 2. Quality control algorithm to assess increasing basal insulin dosage.**
Quality control algorithm to assess increasing basal insulin dosage. User features and
glycaemic outcomes are loaded by the algorithm and assessed for physician-informed
metrics of nocturnal hypoglycaemia, near hypoglycaemia episodes, subject time in target
range, subject adherence, and insulin formulation-dependent requirements.

**Extended Data Fig. 3. Quality control algorithm to assess decreasing basal insulin dosage.**
Quality control algorithm to assess decreasing basal insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for subject adherence, and insulin formulation-dependent requirements.
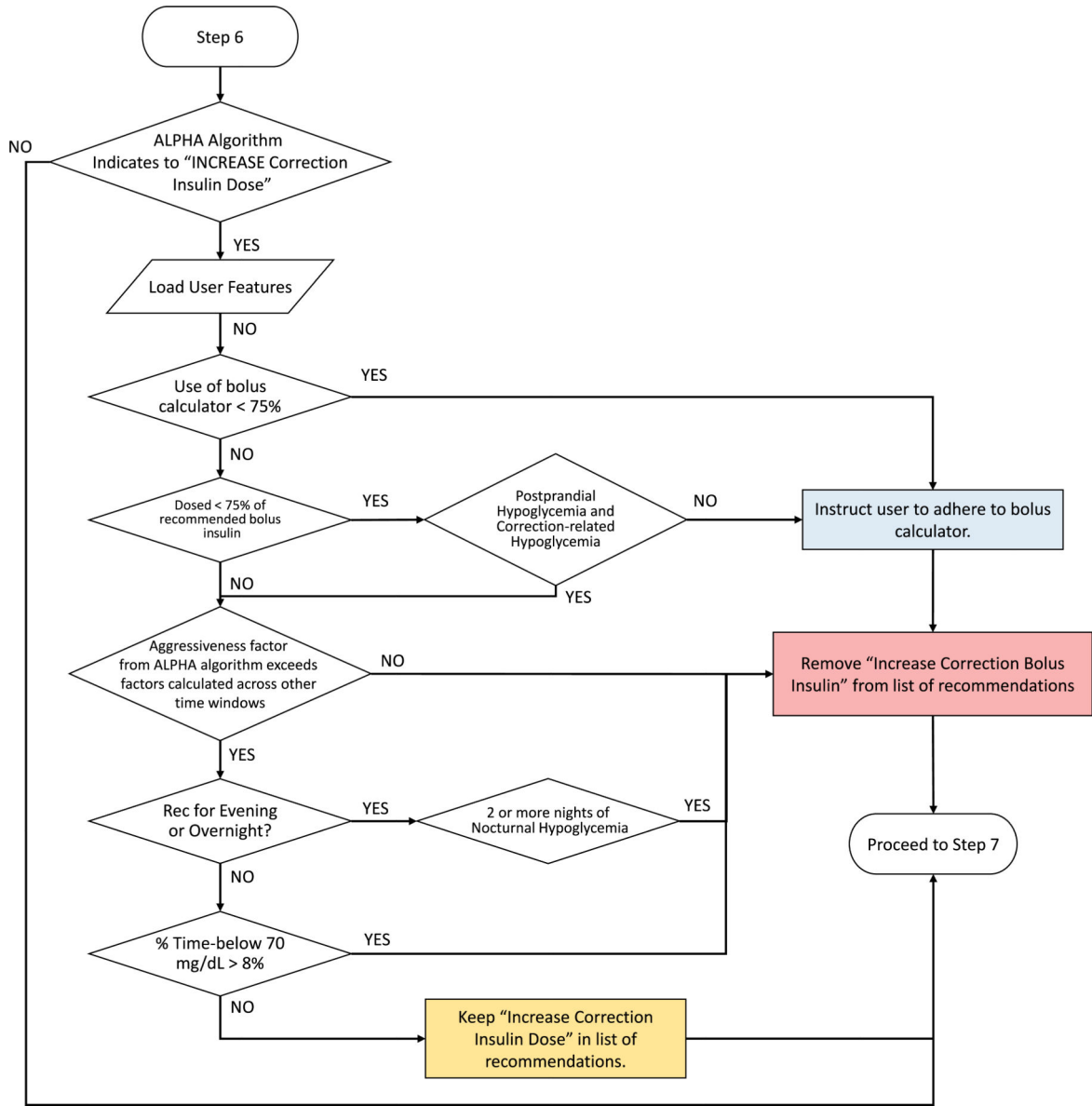
**Extended Data Fig. 4. Quality control algorithm to assess increasing meal bolus insulin dosage.**
Quality control algorithm to assess increasing meal bolus insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of postprandial hypoglycaemia, subject adherence, and factors returned by the ALPHA algorithm.
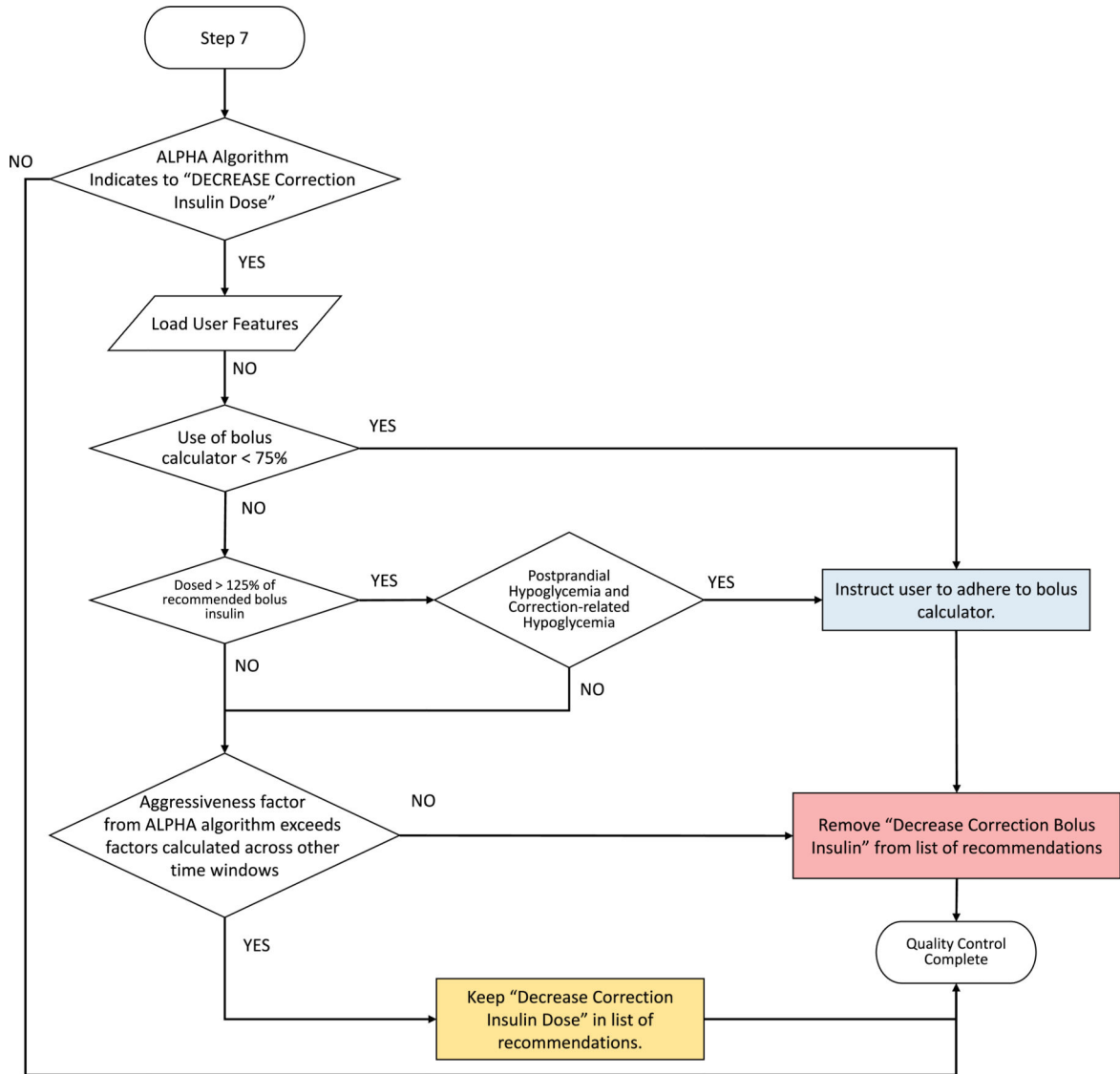
**Extended Data Fig. 5. Quality control algorithm to assess decreasing meal bolus insulin dosage.**
Quality control algorithm to assess decreasing meal bolus insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of postprandial severe hyperglycaemia, subject adherence, and factors returned by the ALPHA algorithm

**Extended Data Fig. 6. Quality control algorithm to assess increasing correction bolus insulin dosage.**

Quality control algorithm to assess increasing correction bolus insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of postprandial and correction-related hypoglycaemia, subject adherence, and factors returned by the ALPHA algorithm.

**Extended Data Fig. 7. Quality control algorithm to assess decreasing correction bolus insulin dosage.**

Quality control algorithm to assess decreasing correction bolus insulin dosage. User features and glycaemic outcomes are loaded by the algorithm and assessed for physician-informed metrics of subject adherence, postprandial and correction-related hypoglycaemia, and factors returned by the ALPHA algorithm.
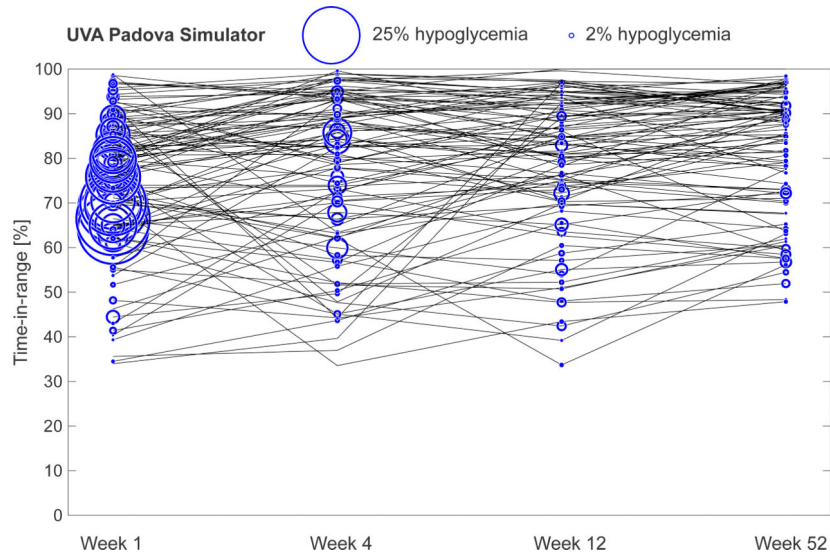
**Extended Data Fig. 8. KNN-DSS engine performance in improving subject outcomes in an independent virtual patient population.**
KNN-DSS engine performance in improving subject outcomes in an independent virtual patient population. Glycaemic outcomes during a 52-week study of the FDA-approved UVA-Padova virtual patient simulator. Percent time in hypoglycaemia is indicated by the blue circular radius.
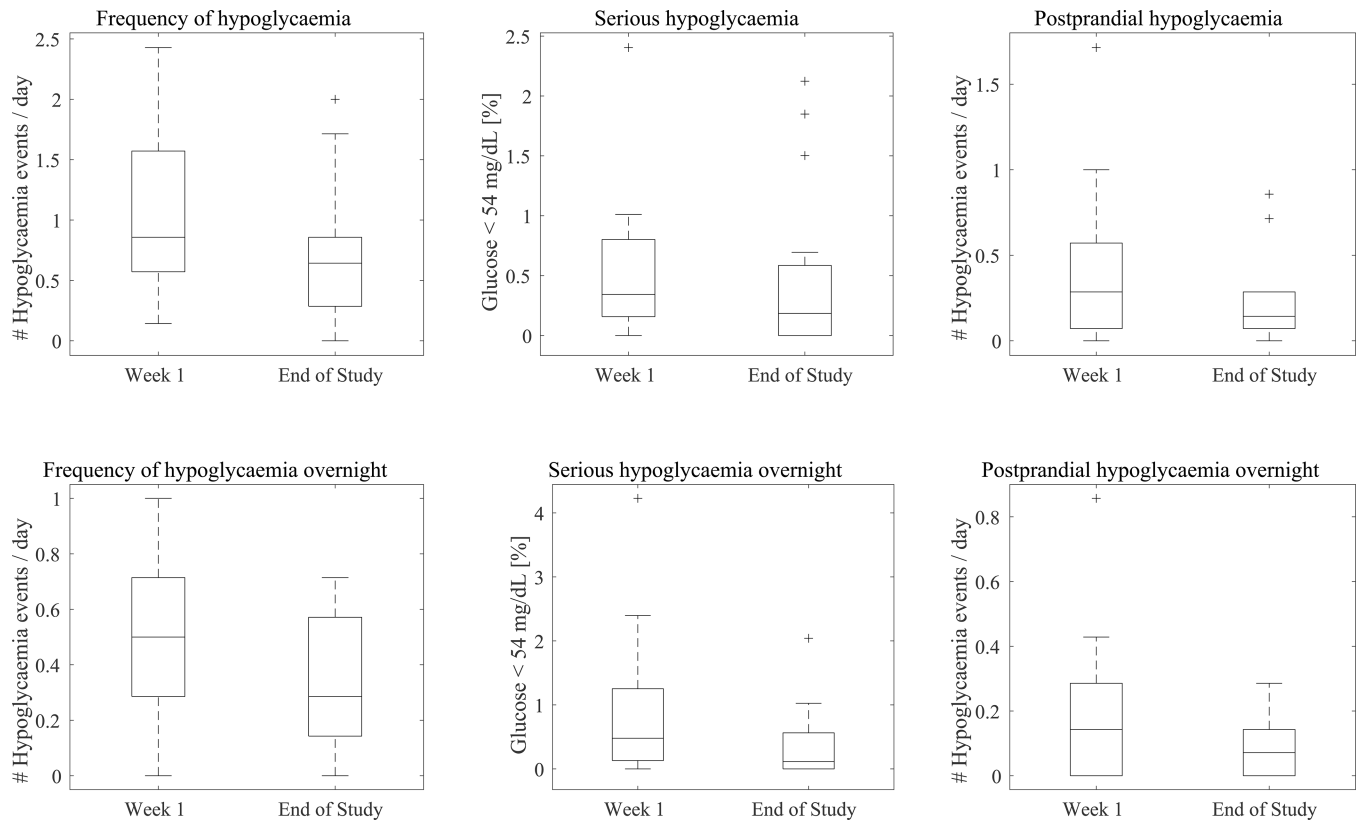
**Extended Data Fig. 9. Outcomes of a human pilot study evaluating KNN-DSS augmented decision support.**

Outcomes of a human pilot study evaluating KNN-DSS augmented decision support over 4 weeks where the first recommendation is given at the start of week 2. For panels a-f, boxplot limits indicate the first and third quartiles, centerline indicates the median, and whiskers mark the last non-outlier data-point within 1.5xIQR. For panels a-f, participant data collected during week 1 and the final week of the study were compared using a two-tailed Wilcoxon signed-rank test, with significance level of 5%. a, Frequency of hypoglycaemia was nominally reduced on the final week compared to week 1 of the study (0.86 vs 0.64, P = 0.051, n = 16 independent subjects). b Serious hypoglycaemia was nominally reduced on the final week compared with week 1 of the study (0.34% vs. 0.19%, P = 0.56, n = 16 independent subjects). c Postprandial hypoglycaemia events were nominally reduced on the final week compared with week 1 (0.29 vs 0.14, P = 0.08, n = 16 independent subjects). d Frequency of overnight hypoglycaemia was significantly reduced on the final week compared to week 1 (0.50 to 0.29, P= 0.04, n = 16 independent subjects). e Serious hypoglycemia overnight was significantly reduced on the final week compared to week 1 (0.48% to 0.11%, P = 0.03, n = 16 independent subjects). f Postprandial hypoglycemia overnight was nominally reduced on the final week compared to week 1 (0.14 to 0.07, P = 0.06, n = 16 independent subjects).

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgements

Competing Interests

The authors declared the following competing of interests regarding research, authorship and publication of this article: J.R.C. and P.G.J. have financial interest in Pacific Diabetes Technologies Inc. (PDT), a company with potential commercial interests in the results and research of this technology. J.R.C. and P.G.J. are founders and shareholders in PDT and P.G.J. is a board member of PDT. Neither J.R.C. nor P.G.J. receive any financial compensation from PDT as consultants or otherwise, beyond the shares in the company. J.R.C. and P.G.J. have received honoraria for consulting and research support from Dexcom. Although the methods on the algorithm were disclosed to the OHSU Technology Transfer Office, there has not yet been a patent filed on the algorithm and PDT does not have any rights to any of the technology described in the paper. N.S.T., C.M.M., R.H.D., L.M.W., D.L.B., V.B.G., F.H.G., W.W.H. and J.E.Y. declare no competing interests.

## References

1. Paris CA et al. Predictors of insulin regimens and impact on outcomes in youth with type 1 diabetes: the SEARCH for Diabetes in Youth study. The Journal of pediatrics 155, 183–189.e181, doi:10.1016/j.jpeds.2009.01.063 (2009). [PubMed: 19394043]

2. Miller KM et al. Current state of type 1 diabetes treatment in the U.S.: updated data from the T1D Exchange clinic registry. Diabetes care 38, 971–978, doi:10.2337/dc15-0078 (2015). [PubMed: 25998289]

3. Resalat N, El Youssef J, Tyler N, Castle J & Jacobs PG A statistical virtual patient population for the glucoregulatory system in type 1 diabetes with integrated exercise model. PloS one 14, e0217301, doi:10.1371/journal.pone.0217301 (2019). [PubMed: 31344037]

4. Cover T & Hart P Nearest neighbor pattern classification. IEEE Trans. Inf. Theor. 13, 21–27, doi:10.1109/tit.1967.1053964 (2006).

5. Nimri R et al. Adjusting insulin doses in patients with type 1 diabetes who use insulin pump and continuous glucose monitoring: Variations among countries and physicians. Diabetes, Obesity and Metabolism 20, 2458–2466, doi:10.1111/dom.13408 (2018).

6. Man CD et al. The UVA/PADOVA Type 1 Diabetes Simulator: New Features. Journal of diabetes science and technology 8, 26–34, doi:10.1177/1932296813514502 (2014). [PubMed: 24876534]

7. Intensive Diabetes Treatment and Cardiovascular Outcomes in Type 1 Diabetes: The DCCT/EDIC Study 30-Year Follow-up. Diabetes care 39, 686–693, doi:10.2337/dc15-1990 (2016). [PubMed: 26861924]

8. Schwartz FL, Guo A, Marling CR & Shubrook JH Analysis of use of an automated bolus calculator reduces fear of hypoglycemia and improves confidence in dosage accuracy in type 1 diabetes mellitus patients treated with multiple daily insulin injections. Journal of diabetes science and technology 6, 150–152, doi:10.1177/193229681200600118 (2012). [PubMed: 22401333]

9. Roze S et al. Cost-effectiveness of continuous subcutaneous insulin infusion versus multiple daily injections of insulin in Type 1 diabetes: a systematic review. Diabetic medicine : a journal of the British Diabetic Association 32, 1415–1424, doi:10.1111/dme.12792 (2015). [PubMed: 25962621]

10. McNally K, Rohan J, Pendley JS, Delamater A & Drotar D Executive Functioning, Treatment Adherence, and Glycemic Control in Children With Type 1 Diabetes. Diabetes care 33, 1159–1162, doi:10.2337/dc09-2116 (2010). [PubMed: 20215458]

11. Sarbacker GB & Urteaga EM Adherence to Insulin Therapy. Diabetes spectrum : a publication of the American Diabetes Association 29, 166–170, doi:10.2337/diaspect.29.3.166 (2016). [PubMed: 27574371]

12. Kirwan M, Vandelanotte C, Fenning A & Duncan JM Diabetes Self-Management Smartphone Application for Adults With Type 1 Diabetes: Randomized Controlled Trial. J Med Internet Res 15, e235, doi:10.2196/jmir.2588 (2013). [PubMed: 24225149]

13. Charpentier G et al. The Diabeo software enabling individualized insulin dose adjustments combined with telemedicine support improves HbA1c in poorly controlled type 1 diabetic patients: a 6-month, randomized, open-label, parallel-group, multicenter trial (TeleDiab 1 Study). Diabetes care 34, 533–539, doi:10.2337/dc10-1259 (2011). [PubMed: 21266648]

14. Wu Y et al. Mobile App-Based Interventions to Support Diabetes Self-Management: A Systematic Review of Randomized Controlled Trials to Identify Functions Associated with Glycemic Efficacy. JMIR mHealth and uHealth 5, e35, doi:10.2196/mhealth.6522 (2017). [PubMed: 28292740]

15. Veazie S et al. Rapid Evidence Review of Mobile Applications for Self-management of Diabetes. Journal of general internal medicine, doi:10.1007/s11606-018-4410-1 (2018).

16. Beck RW et al. Effect of Continuous Glucose Monitoring on Glycemic Control in Adults With Type 1 Diabetes Using Insulin Injections: The DIAMOND Randomized Clinical Trial. Jama 317, 371–378, doi:10.1001/jama.2016.19975 (2017). [PubMed: 28118453]

17. Steil GM et al. Use of Automated Clinical Decision Support (CDS) to Effect Glycemic Control in Elderly Patients with T1D. Diabetes 67, 921–P, doi:10.2337/db18-921-P (2018).

18. Palerm CC, Zisser H, Jovanovic L & Doyle FJ 3rd. A Run-to-Run Control Strategy to Adjust Basal Insulin Infusion Rates in Type 1 Diabetes. J Process Control 18, 258–265, doi:10.1016/j.jprocont.2007.07.010 (2008). [PubMed: 18709180]

19. Herrero P, Bondia J, Gimenez M, Oliver N & Georgiou P Automatic Adaptation of Basal Insulin Using Sensor-Augmented Pump Therapy. Journal of diabetes science and technology 12, 282–294, doi:10.1177/1932296818761752 (2018). [PubMed: 29493359]

20. Toffanin C, Messori M, Cobelli C & Magni L Automatic adaptation of basal therapy for Type 1 diabetic patients: A Run-to-Run approach. Biomedical Signal Processing and Control 31, 539–549, doi:10.1016/j.bspc.2016.09.002 (2017).

21. Zisser H, Palerm CC, Bevier WC, Doyle FJ 3rd & Jovanovic L Clinical update on optimal prandial insulin dosing using a refined run-to-run control algorithm. Journal of diabetes science and technology 3, 487–491, doi:10.1177/193229680900300312 (2009). [PubMed: 20144286]

22. Herrero P et al. Advanced Insulin Bolus Advisor Based on Run-To-Run Control and Case-Based Reasoning. IEEE journal of biomedical and health informatics 19, 1087–1096, doi:10.1109/jbhi.2014.2331896 (2015). [PubMed: 24956470]

23. Perez-Gandia C et al. Decision Support in Diabetes Care: The Challenge of Supporting Patients in Their Daily Living Using a Mobile Glucose Predictor. Journal of diabetes science and technology 12, 243–250, doi:10.1177/1932296818761457 (2018). [PubMed: 29493361]

24. Breton MD et al. Continuous Glucose Monitoring and Insulin Informed Advisory System with Automated Titration and Dosing of Insulin Reduces Glucose Variability in Type 1 Diabetes Mellitus. Diabetes technology & therapeutics 20, 531–540, doi:10.1089/dia.2018.0079 (2018). [PubMed: 29979618]

25. Reddy M et al. Clinical Safety and Feasibility of the Advanced Bolus Calculator for Type 1 Diabetes Based on Case-Based Reasoning: A 6-Week Nonrandomized Single-Arm Pilot Study. Diabetes technology & therapeutics 18, 487–493, doi:10.1089/dia.2015.0413 (2016). [PubMed: 27196358]

26. Resalat N, El Youssef J, Reddy R, Castle J & Jacobs PG Adaptive tuning of basal and bolus insulin to reduce postprandial hypoglycemia in a hybrid artificial pancreas. Journal of Process Control 80, 247–254, doi:10.1016/j.jprocont.2019.05.018 (2019).

27. Sørensen TJ A method of establishing groups of equal amplitude in plant sociology based on similarity of species content and its application to analyses of the vegetation on Danish commons. (I kommission hos E. Munksgaard, 1948).

28. Davidson MB, Duran P, Davidson SJ & Lee M Comparison of Insulin Dose Adjustments by Primary Care Physicians and Endocrinologists. Clinical Diabetes 36, 39, doi:10.2337/cd17-0021 (2018). [PubMed: 29382977]

29. Bashan E & Hodish I Frequent insulin dosage adjustments based on glucose readings alone are sufficient for a safe and effective therapy. Journal of Diabetes and its Complications 26, 230–236, doi:10.1016/j.jdiacomp.2012.03.012 (2012). [PubMed: 22521317]

30. Reddy R et al. The effect of exercise on sleep in adults with type 1 diabetes. Diabetes, obesity & metabolism 20, 443–447, doi:10.1111/dom.13065 (2018).

## References (Methods Only)

31. Castle JR et al. Randomized Outpatient Trial of Single- and Dual-Hormone Closed-Loop Systems That Adapt to Exercise Using Wearable Sensors. Diabetes care 41, 1471–1477, doi:10.2337/dc18-0228 (2018). [PubMed: 29752345]

32. Pettus J & Edelman SV Recommendations for Using Real-Time Continuous Glucose Monitoring (rtCGM) Data for Insulin Adjustments in Type 1 Diabetes. Journal of diabetes science and technology 11, 138–147, doi:10.1177/1932296816663747 (2017). [PubMed: 27530720]

33. Whitney AW A Direct Method of Nonparametric Measurement Selection. IEEE Trans. Comput. 20, 1100–1103, doi:10.1109/t-c.1971.223410 (1971).

34. Scheiner G Practical CGM: improving patient outcomes through continuous glucose monitoring. (American Diabetes Association, 2015).

35. White JW, Rassweiler A, Samhouri JF, Stier AC & White C Ecologists should not use statistical significance tests to interpret simulation model results. Oikos 123, 385–388, doi:10.1111/j.1600-0706.2013.01073.x (2014).
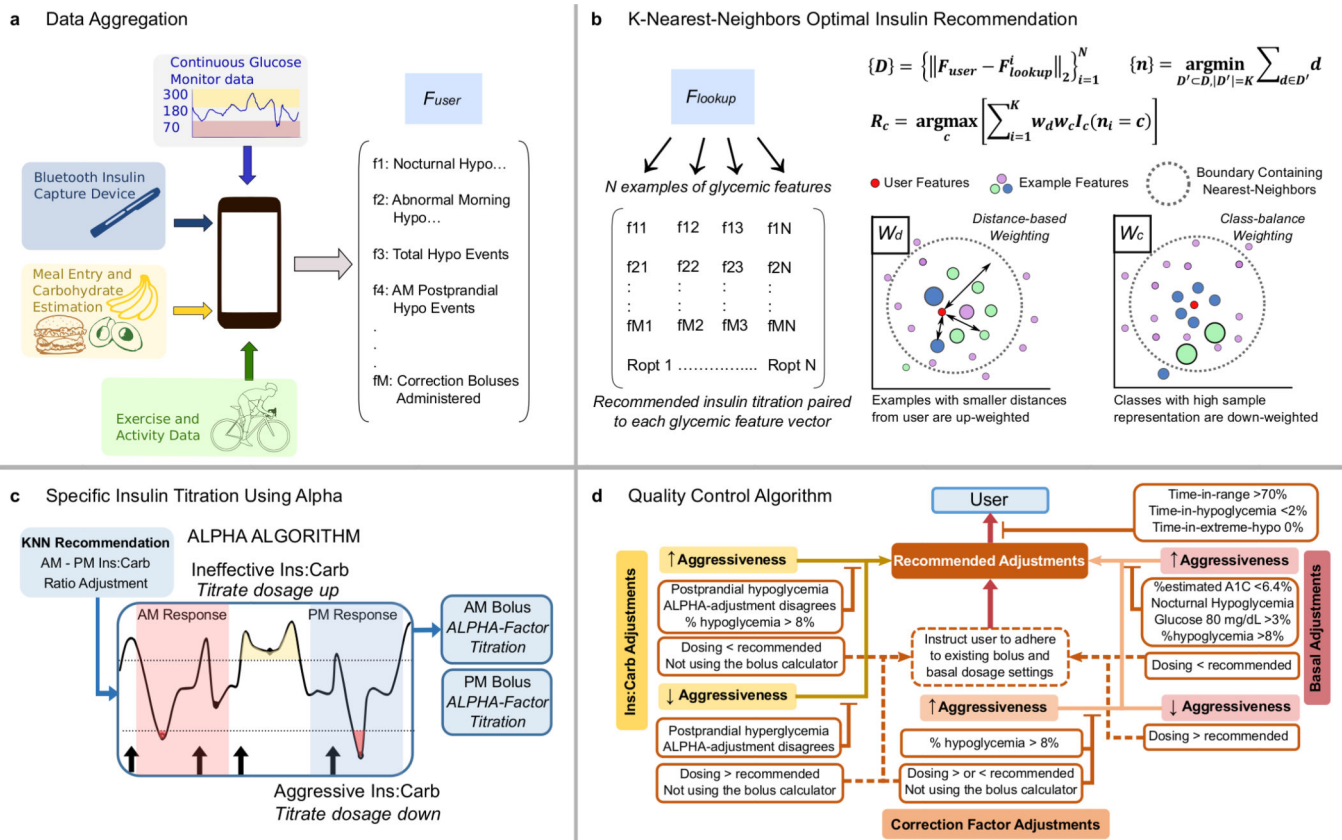
**Figure 1: Decision support engine framework to identify user-specific insulin titrations.**
**a**, The user data are aggregated and processed for extracting glucose, insulin, meal, and exercise features that may be used to optimally titrate insulin doses. **b**, The user features $F$_user are matched to the closest examples in the look-up table for the K-nearest neighbours algorithm, $F$_lookup. The distance between user features and the examples in the look-up table are calculated as $\{D\}$, and the $K$ examples within minimum distance to user features, $\{n\}$, are weighted by distance, $w$_d, and class-size, $w$_c; the final insulin dosage recommendations, $R$_c, are returned by the KNN algorithm. **c**, For those recommendations indicated by the KNN-DSS, the ALPHA algorithm assigns an aggressiveness factor that titrates carbohydrate ratios and correction factors to improve time in target range and reduce time in hypoglycaemia. **d**, A quality control algorithm is employed to ensure that KNN-DSS recommendations adhere to physician standards
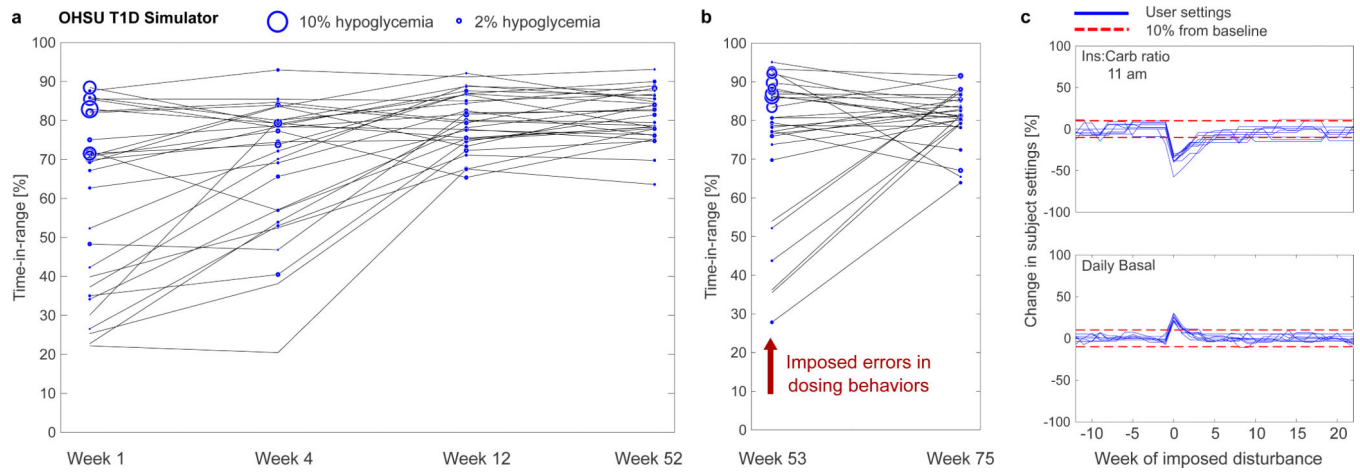
**Figure 2: Engine performance in improving subject outcomes *in silico*.**
**a**, Outcomes of the *in silico* evaluation of the KNN-DSS over 52 weeks. Virtual patients from the OHSU T1D simulator undergo weekly use of the KNN-DSS engine. **b,** At 52 weeks, new insulin settings and dosing behaviors are imposed on patients, the effects of which are measured at 53 weeks. **c,** Evolution of patient settings 20 weeks from baseline following a disturbance at week 52. The red dashed line indicates 10% from baseline.

**Table 1|**

Recommendations delivered by the KNN-DSS engine.

| Recommendation | Message to User | Adjustment Window | Titration Method |
|---|---|---|---|
| Basal Adjustment | "You may need (less/more) basal insulin. It is recommended that you (decrease/increase) your AAA insulin from BBB to CCC units. | AM<br>PM | *Adjustment by +/− 10% from weekly baseline settings* |
| Carb:ins Ratio | "You may need (less/more) insulin before (breakfast/lunch/dinner/a meal). It is recommended that you change your carb ratio from AAA to BBB." | 7:00 – 11:00<br>11:00 – 15:00<br>15:00 – 20:00<br>20:00 – 7:00 | *ALPHA Algorithm: meal bolus glycaemic response and assignment of dosage titration.* |
| Correction Factor | "You may need (less/more) insulin in (morning / afternoon / evening / nighttime) to bring down high glucose levels. It is recommended that you change your correction factor from AAA to BBB." | 7:00 – 11:00<br>11:00 – 15:00<br>15:00 – 20:00<br>20:00 – 7:00 | *ALPHA Algorithm: correction bolus glycaemic response and assignment of dosage titration.* |
| Bolus Adherence[a] | "You may have taken (more/less) insulin than recommended by the bolus calculator during certain times of day. It is recommended that you use the amount recommended by the bolus calculator" | All Day | N/A |
| Basal Adherence[a] | "We have found that the amount of basal insulin that you are taking is different than the amount we recommend. Taking the recommended amount may improve your glucose levels." | All Day | N/A |

Recommendations for insulin dosage are titrated to be higher or lower during different time windows using the specified titration method.

[a]Behavioral recommendation.

**Table 2|**

Agreement between KNN-DSS and endocrinologist recommendations using real-world human data.

| Recommendation Comparison | % Agreement | % Disagreement | % Additional | % Not comparable |
|---|---|---|---|---|
| **Assessing recommendations on all participant data (N = 78 weeks)** | *Full* 27.8 <br> *Partial* 40.1 <br> Overall **67.9** | *Full* 0.4 <br> *Partial* 6.0 <br> Overall **6.4** | *Physician* 6.4 <br> *Engine* 10.3 <br> Overall **16.7** | **9.0** |
| **Assessing recommendations on participant data with > 75% recommendation adherence (N = 39 weeks)** | *Full* 50.8 <br> *Partial* 16.7 <br> Overall **67.5** | *Full* 2.6 <br> *Partial* 8.8 <br> Overall **11.4** | *Physician* 13.2 <br> *Engine* 7.9 <br> Overall **21.1** | **0.0** |

Agreement is calculated using a Sorensen-Dice coefficient similarity comparing physician recommendations to engine recommendations for each week of observed data. Adherence refers to participant use of the bolus calculator.