Protocol and statistical analysis plan supplements for the randomized clinical trial assessing efficacy and safety of an Artificial-Pancreas-like Closed-loop control in adults with type 2 diabetes on Multiple Daily Injections therapy (APC MID)

This supplement contains the following items:
1. Original protocol, final protocol, summary of changes.
2. Original statistical analysis plan, final statistical analysis plan, summary of changes.

APC MID – RCT assessing efficacy and safety of an artificial-pancreas-like closed-loop control in adults with type 2 diabetes on multiple daily injections therapy

Protocol Summary

IRB Approval Date:	15/Jan/2021
Peking University People's Hospital IRB #:	2020PHB338-01
Sponsor:	Peking University People's Hospital Scientific Research Development Funds, Peking University Medical Science and Technology Fund, National Natural Science Foundation of China, BIT Research and Innovation Promoting Project
Principal Investigators:	Wei Liu Linong Ji
Internal Staff and Sub-Investigators:	Rui Zhang Xiuying Zhang Xiaoling Cai Xueyao Han Yingying Luo Meng Li Deheng Cai Wenjing Wu Yumin Ma Dawei Shi

Background and Introduction

Diabetes Mellitus (DM) is a disease that causes related tissues and organs to malfunction due to homeostasis disorder of glucose metabolism and increased blood sugar levels, and the pathophysiology of DM is caused by insufficient or decreased insulin secretion. Epidemiological studies in recent years have shown that the incidence of diabetes is increasing worldwide. The latest statistics show that the current global diabetes patients are 415 million, and about 1/4 of them are from China. In addition to the impact on the family and society, diabetes has also caused a heavy burden on China's economic development, and it is estimated that the annual direct economic cost due to diabetes is about 91 billion RMB [1].

As a chronic noncommunicable disease, maintaining long-term adequate blood glucose control can greatly improve the prognosis and reduce the burden on individuals and society. At present, in diabetes treatment, insulin is still the "ultimate weapon" to control blood sugar, based on real-world research which shows that about 1/4 of diabetic patients are currently receiving insulin therapy [2]. However, despite continuous improvements in devices and drugs over the past 100 years since the invention of insulin, the average glycated hemoglobin (Hemoglobin A1, HbA1c) of insulin users is still as high as 8.5%, with only about 1/3 reaching the guideline recommended range of less than 7.0%, and 1/3 exceeding 9.0% [3,4].

A growing number of studies have shown that continuous adjustment of insulin injection doses according to inter- and intra-individual differences can improve the effectiveness of insulin therapy in blood glucose control [5]. However, in clinical practice, due to the limitations of consultation time, resources, and patients' ability to receive relevant training, the average dose adjustment cycle of patients receiving insulin therapy is 3-6 months. At the same time, a study in type 2 diabetic patients receiving insulin therapy showed that insulin therapists often failed to use insulin doses that satisfactorily controlled their blood glucose [6]. How to adjust the dose of adjuvant insulin therapy more effectively and safely and promote the translation of guideline opinions into clinical practice is an urgent problem in the field of diabetes treatment.

In type 1 diabetes, the "artificial pancreas," or closed-loop insulin pump, can hopefully bridge the gap between guidelines and reality. It monitors blood glucose in real time and enables precise dose adjustment of insulin injections, and related studies have also shown its advantages in blood glucose control [7]. A randomized controlled study by Thabit et al. in children and adults with type 1 diabetes showed that closed-loop insulin pumps were able to reduce HbA1c from 8.5% to

7.5% without increasing the occurrence of hypoglycemia [8]. However, the high cost of closed-loop insulin pumps limits their use in the larger population of type 2 diabetes.

Some self-glucose monitoring devices can upload blood glucose values to medical staff in the form of remote data transmission to help patients adjust insulin doses under the guidance of medical professionals [9]. Some devices then accept blood glucose data and give recommendations for adjusting the dose of insulin according to a preset algorithm, which is provided to medical staff, who then pass on to patients through voice or video communication [10]. However, these tools still depend on the involvement of healthcare providers, who have very limited time to allocate to these aspects in practice. At the same time, there are mobile phone apps that digitize insulin dose adjustment methods to help patients adjust insulin doses; The apps are approved by the US Food and Drug Administration (USFDA), require a clinician's prescription to obtain them [11]; although these apps have long existed, diabetics treated with insulin have not seen improvement in recent years [3].

In recent years, the rapid development of bioinformatics and big data science has brought new opportunities and means for disease prevention, diagnosis and treatment. The d-Nav insulin Guidance System (Insulin Guidance System) developed by Hygieia in the United States was officially launched in the UK in 2012, which can provide guidance and suggestions for the dose of insulin injection by analyzing the results of self-blood glucose monitoring data and previous insulin injection doses, combined with blood glucose values before insulin injection. The randomized controlled clinical study [2] published by Lancet in February 2019 in people with type 2 diabetes showed that HbA1c levels decreased by 1% from baseline 8.7% in the intervention group after 6 months of using the d-Nav insulin injection dose guidance system supplemented by physician guidance, and the control group reduced HbA1c by 0.3% from 8.5% after only 6 months under physician guidance ($p<0.0001$). This suggested that the d-Nav insulin injection dose guidance system supplemented with physician-guided glycemic control was significantly better than physician-guided alone.

We searched for domestic scholars' research in this field, and although there have been some preliminary explorations in closed-loop insulin pumps, the field of providing dose adjustment recommendations through artificial intelligence algorithms based on blood glucose monitoring data is still open/blank. Pioneering research in this field can bridge the gap between the domestic

and international fields and lay the foundation for the development of related equipment in the future.

In the first part of this study, an AI-based mealtime insulin dose adjustment algorithm has been established, and the recommendation of algorithm and the clinician's actual decision have been compared in patients with type 2 diabetes who received insulin injections, based on which the algorithm was further improved.

In our preliminary study, an individualized decision-making system for pre-prandial insulin dose based on Gaussian process was established through effectively exploiting the historical data information of diabetic patients and establishing a postprandial blood glucose prediction model, based on which individualized insulin dosage decision was achieved

using risk-sensitive control ideas for improved postprandial blood glucose levels while ensuring safety. The basic methodological idea is described as follows:

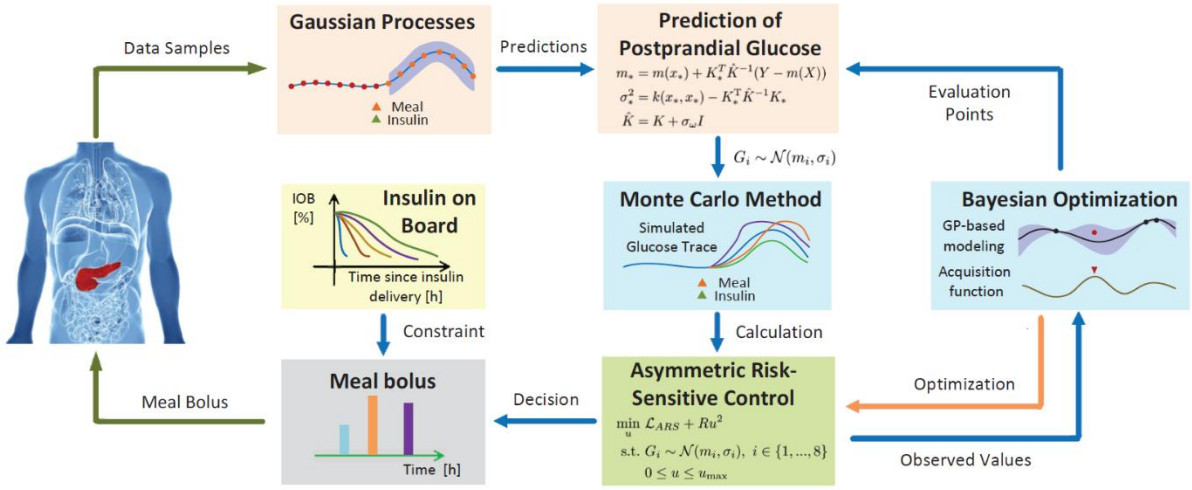


Figure 1 Mealtime insulin dose adjustment algorithm and optimization design ideas

The overall structure of the proposed data-driven meal bolus decision method is illustrated in Figure 1. The method builds on three key components: model learning, asymmetric risk-sensitive control, and Bayesian optimization. Model learning is responsible for constructing the postprandial glucose dynamics. Here, the aim is to provide a robust description for postprandial glucose dynamics using Gaussian Processes (GPs). Specifically, fed with the serialized data samples (including pre-prandial glucose measurements, the corresponding meal information, bolus dosage and postprandial glucose measurements), the GPs are trained and then applied to provide the prediction and the uncertainty estimation of postprandial glucose trajectories.

Considering the asymmetric risks of hyper- and hypoglycemia and the uncertainties in the predicted glucose trajectories, we develop an asymmetric risk-sensitive cost function to favor safe control actions. Finally, a constrained stochastic optimization problem is formulated for the meal bolus decision based on the designed cost function. Since the gradient of the cost function is unavailable, we solve the optimization problem using Bayesian optimization and Monte-Carlo simulations. To ensure the safety of the method, insulin on board (IOB) constraints are also incorporated.

Model learning: The postprandial glucose (PG) dynamics is described in a multistep form using autoregressive models, and each step is separately represented by a nonlinear function with additive noise, which has the form:

$$\begin{aligned} P_{t+i} &= f_{t+i}(\boldsymbol{z}_{t+i-1}) + w_{t+i}, \\ \boldsymbol{z}_{t+i-1} &= [P_{t+i-1-l}, \cdots, P_{t+i-1}, u, d], \end{aligned} \tag{1}$$

where t denotes the time of the meal intake, $i = \{1, 2, \cdots, n\}$ represents the time step after the meal intake, w is a white Gaussian noise, P is the glucose measurement, and l is the lag for autoregressive outputs; u is the meal bolus, and d is the carbohydrate intake. To convey the most information of glucose situations, the lag of l = 7 and the sampling period of T = 15 min are considered. This corresponds to the lag of 2 hours. Correspondingly, we take n = 8; since the sampling period is 15 min, this corresponds to the duration of 2 hours.

The utilization of GPs is divided into two stages: modeling and prediction. In the modeling stage, based on equation (1), the GPs are separately used to model the PG dynamics in each step following a similar way. For example, for the time step $t + 1$, we use $\boldsymbol{z}_t = [P_{t-7}, \cdots, P_t, u, d]^{\mathrm{T}}$ as training inputs, and the differences $\Delta P_t = P_{t+1} - P_t$ as training targets to reduce the prediction uncertainty. To incorporate prior knowledge into the GP, a linear mean function is adopted for the GP. Moreover, a commonly-used covariance function known as squared exponential (SE) covariance kernel is considered. Note that u and d stay the same for all steps. To highlight the effects of u and d on the glucose regulation, the glucose measurements in the training inputs for each step are separately normalized into [0, 1] using min-max normalization.

In the prediction stage, given a bolus dosage and known carbohydrate amount intakes, by iteratively feeding back the predicted mean of previous step into the input for the prediction, we can obtain the prediction for each step using corresponding trained GPs. This prediction corresponds to the difference between the current step and previous step. We then add this

prediction with the predicted mean of the previous step to determine the final prediction in current step. Uniting the predictions for the all steps, the GPs are able to provide 8-step predictions for the PG trajectories in respond to pre-prandial glucose situations, carbohydrate intakes and meal boluses. Note that when the eating habit of a subject is approximately consistent in terms of timing and sizes of meal intake, we can approximate the effect of similar food intake as an invariant disturbance and the meal size information can be optional for postprandial glucose prediction, utilizing the robust prediction ability of the GPs; this allows the design of a bolus decision algorithm without meal announcements.

Asymmetric risk-sensitive control: We denote the 8-step predictions provided by the GPs as $G_i \sim N(m_i, \sigma_i^2)$, $i = \{1, 2, \cdots, 8\}$, respectively, and collect them as a vector state $\boldsymbol{G} = [G_1, G_2, \cdots, G_8]^{\mathrm{T}}$, which describes the probability distribution of postprandial glucose trajectories. According to the principle of risk-sensitive analysis, an asymmetric risk-sensitive (RS) cost is designed as follows:

$$L_{ARS} = -\frac{2}{\gamma}\log\mathbb{E}\left[\exp\left(-\frac{1}{2\gamma}\left((\boldsymbol{G}-\boldsymbol{G_r})_+^{\mathrm{T}}\boldsymbol{Q}^+(\boldsymbol{G}-\boldsymbol{G_r})_+ + (\boldsymbol{G}-\boldsymbol{G_r})_-^{\mathrm{T}}\boldsymbol{Q}^-(\boldsymbol{G}-\boldsymbol{G_r})_-\right)\right)\right], \quad (2)$$

where

$$(\boldsymbol{G}-\boldsymbol{G_r})_+ = [(G_1 - G_{r1})\mathbf{1}(G_1 - G_{r1} \geq 0), \ldots, (G_8 - G_{r8})\mathbf{1}(G_8 - G_{r8} \geq 0)]^{\mathrm{T}}, \quad (3)$$

$$(\boldsymbol{G}-\boldsymbol{G_r})_- = [(G_1 - G_{r1})\mathbf{1}(G_1 - G_{r1} < 0), \ldots, (G_8 - G_{r8})\mathbf{1}(G_8 - G_{r8} < 0)]^{\mathrm{T}}, \quad (4)$$

and $\mathbf{1}(\cdot)$ denotes the indicator function; $\boldsymbol{G_r}$ is the target for the postprandial glucose management; $\boldsymbol{Q}^+$ is a positive penalty matrix for the glucose excursions above the target; $\boldsymbol{Q}^-$ is a negative penalty matrix for the glucose excursions below the target; $\gamma < 0$ is a risk sensitivity parameter that determines the system's attitude towards uncertainty. As for the design of asymmetric penalty, $\boldsymbol{Q}^+$ is designed as a constant diagonal matrix. Based on the designed $\boldsymbol{Q}^+$, the diagonal elements of $\boldsymbol{Q}^-$ are devised to increase exponentially with the increase of the absolute deviation from target while being restricted by upper and lower bounds. Specifically, the ith diagonal element of $\boldsymbol{Q}^-$ has the form of

$$Q_i^- := Q_i^+(c_1/(1+\exp\{\alpha(\beta - |G_i - G_{ri}|)\}) + c_2), \quad (5)$$

where Q_i^+ is the ith diagonal element of $\boldsymbol{Q}^+$; $\Gamma := [\alpha, \beta, c_1, c_1]$ is a quadruple determines the penalty intensity, which is designed same for all diagonal elements. The lower and upper bounds are determined by c_2 and $c_1 + c_1$, respectively, and the rate of increase is controlled by α. Given the designed asymmetric risk-sensitive cost L_{ARS} in (2), a GP-based asymmetric risk-

sensitive control for the meal bolus decision is formulated as the following constrained stochastic optimization problem:

$$\min_{u} L(u) \coloneqq L_{ARS} + R\big(u - u_p\big), \tag{6}$$

$$\text{s.t. } G_i \sim N\big(m_i, \sigma_i^2\big), i = \{1,2,\cdots,8\}, \tag{7}$$

$$-3 \le u - u_p \le 3, \tag{8}$$

$$0 \le u \le u_{\max}, \tag{9}$$

where u is the meal bolus to be optimized, u_p denotes the meal bolus for the previous same meal type, R is the input weight to compromise the asymmetric RS cost and the actual needed bolus dosage, m_i and σ_i^2 are parameters provided by the GPs.

Bayesian optimization: Since the gradient of the cost function cannot be obtained analytically, Bayesian optimization (BO) is employed to solve the above constrained stochastic optimization problem. Fundamentally, the Bayesian optimization is a sequential query-based approach to solving an optimization problem. Through building a probabilistic surrogate model that consists of a prior distribution that captures the behavior of the unknown objective function, an acquisition function is developed to determine the optimal sequence of queries. After observing the output of the objective function at each query, the prior is updated to produce a more informative posterior distribution that describes the behavior of objective function.

Here, we utilize a new GP to construct an approximation of a complex map from the decision variable u to the cost function value $L(u)$ in (6). Specifically, we consider a prior zero mean function and the SE covariance function with a scalar input. Given the set of N past observations $D_{1:N} = \{u_{1:N}, L(u_{1:N})\}$, the GP is trained and then applied to predict the cost function value for a candidate meal bolus u_*. The prediction value is denoted as $\hat{L}(u_*)$ and will be utilized to construct the acquisition function. Note that the value of the cost function corresponding to the decision variable is estimated by Monte-Carlo simulations. Specifically, given a bolus dosage, we generate 1000 samples for the postprandial glucose trajectory based on the joint Gaussian distribution provided by the GP. The average cost for these samples is further calculated and regarded as the observed value of the cost corresponding to the bolus dosage.

As a critical ingredient of the BO, the acquisition function guides the optimization by determining the optimum candidate point for the next evaluation. Specifically, utilizing the

prediction information offered by the model learning phase, the acquisition function is constructed to determine the candidate point by maintain a trade-off exploration of the search space and exploitation of current promising areas. Here, the expected improvement acquisition function is considered. After M sequential operations of the BO, the final solution u_b is determined. For safety concern, an IOB constraint is enforced to prevent over-bolus based on insulin delivery history. Denoting the IOB constraint as u_{IOB}, the final meal bolus is determined as $\widetilde{u}_b = u_b - u_{IOB}$.

The performance of the proposed algorithm was evaluated on the 10-adult cohort of the US Food and Drug Administration (FDA) accepted Universities of Virginia (UVA)/Padova T1DM simulator, and compared with the standard meal bolus calculator. For the case of announced meals, the proposed method achieved satisfactory and similar performance for scenarios of nominal basal rates, in terms of mean glucose level and percent time in the safe range, without increasing the risk of hypoglycemia. Similar results were observed for the case without the meal information (assuming that the patient follows a consistent diet). Finally, the comparison performance in terms of glucose regulation and meal bolus were further illustrated in Figures 2 and 3, respectively, where the 5%, 25%, 75% and 95% quartile curves together with the median curves were presented.

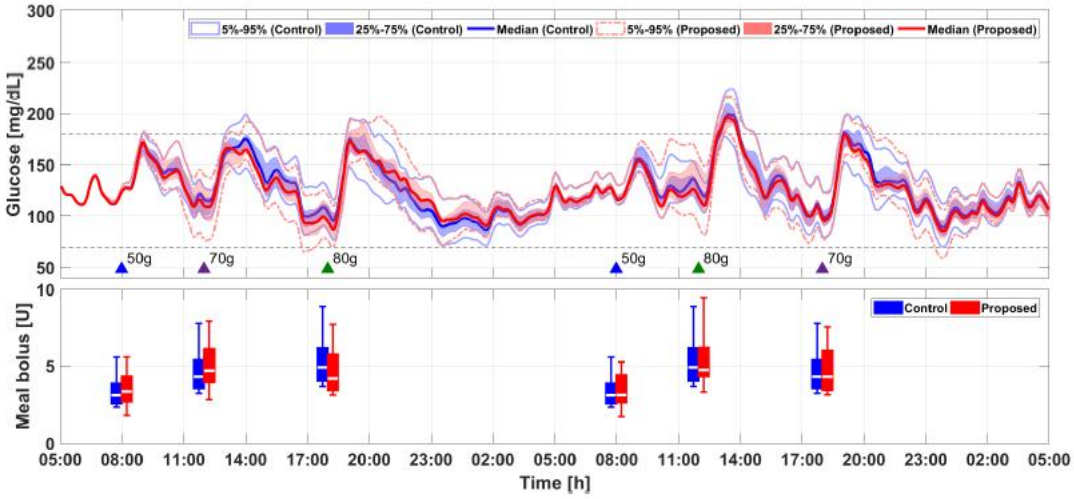


Figure 2 Performance comparison with meal information in terms of glucose regulation and meal bolus (red and blue plots denote the proposed algorithm and standard meal bolus calculator, respectively). Yellow, blue, green and purple triangles denote meals of 45 g, 55 g, 65 g and 85 g CHO, respectively.

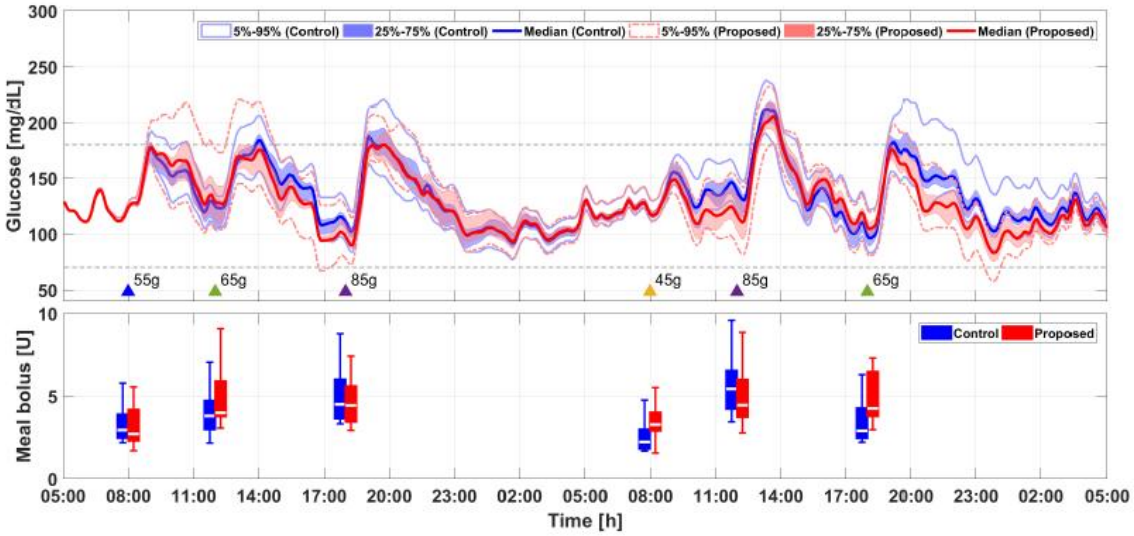


Figure 3 Performance comparison without meal information in terms of glucose regulation and meal bolus (red and blue plots denote the proposed algorithm and standard meal bolus calculator, respectively). Yellow, blue, green and purple triangles denote meals of 45 g, 55 g, 65 g and 85 g CHO, respectively.

Additionally, advisory mode analysis based on clinical data was also performed to evaluate the safety and effectiveness of the proposed approach on realistic clinical application scenarios. Specifically, the advisory-mode analysis allows the comparisons with insulin recommendations made by clinicians, through feeding the identical glucose data obtained in the clinical trial to the system. Here, by feeding the historical pre-prandial glucose data to the proposed method, the corresponding meal boluses are determined and compared with the ones following the clinician's advice. The obtained meal boluses have no causal impact on historical data and are only for the comparison. The results showed that the proposed method could determine similar meal boluses compared with the ones determined by the physicians. The results for a particular comparison were presented in Figure 4.

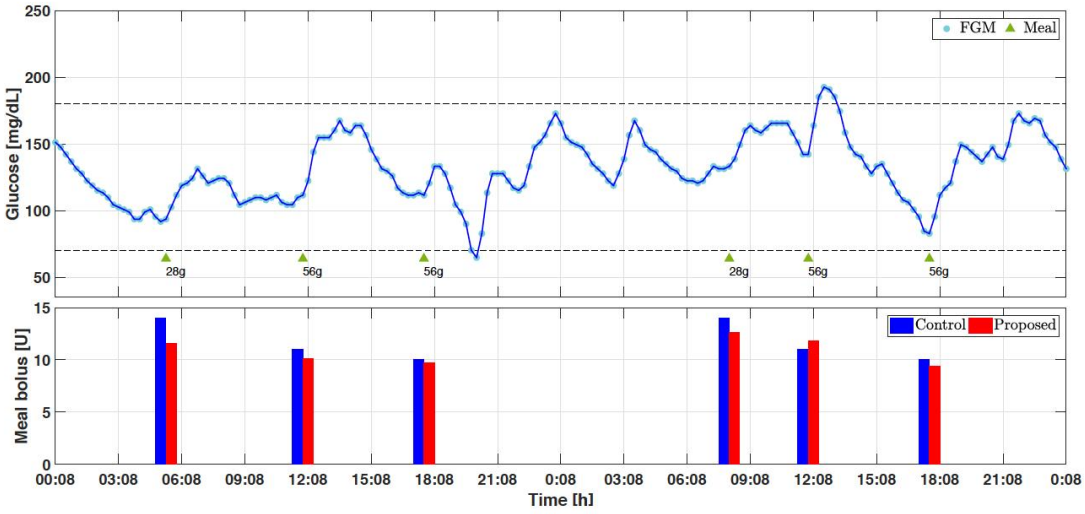


Figure 4 Performance evaluation of the proposed method based on the clinical data (red and blue bars denote insulin dosage recommended by the proposed method and physicians, respectively). Meals are denoted by green triangles with sizes below them.

References

1. IDF Diabetes Atlas, 2017, 8th Edition, http://www.diabetesatlas.org

2. Bergenstal RM, Johnson M, Passi R, et al. Automated insulin dosing guidance to optimise insulin management in patients with type 2 diabetes: a multicentre, randomised controlled trial. Lancet 2019;393:1138-1148.

3. Selvin E, Parrinello CM, Daya N, Bergenstal RM. Trends in insulin use and diabetes control in the U.S.: 1988–1994 and 1999–2012. Diabetes Care 2015; 39: e33–35.

4. Grabner M, Chen Y, Nguyen M, Abbott SD, Quimbo R. Using observational data to inform the design of a prospective effectiveness study for a novel insulin delivery device. Clinicoecon Outcomes Res 2013; 5: 471–79.

5. Davidson MB. How our current medical care system fails people with diabetes: lack of timely, appropriate clinical decisions. Diabetes Care 2009; 32: 370–72.

6. Rosenstock J, Ahmann AJ, Colon G, Scism-Bacon J, Jiang H, Martin S. Advancing insulin therapy in type 2 diabetes previously treated with glargine plus oral agents: prandial premixed (insulin lispro protamine suspension/lispro) versus basal/bolus (glargine/lispro) therapy. Diabetes Care 2008; 31: 20–25.

7. Doyle FJ 3rd, Huyett LM, Lee JB, et al. Closed-loop artificial pancreas systems: engineering the algorithms. Diabetes Care 2014;37:1191-1197.
8. Thabit H, Tauschmann M, Allen JM, et al. Home use of an artificial beta cell in type 1 diabetes. N Engl J Med 2015; 373: 2129–2140.
9. Mora P, Buskirk A, Lyden M, Parkin CG, Borsa L, Petersen B. Use of a novel, remotely connected diabetes management system is associated with increased treatment satisfaction, reduced diabetes distress, and improved glycemic control in individuals with insulin-treated diabetes: first results from the personal diabetes management study. Diabetes Technol Ther 2017; 19: 715–22.
10. Padman R, Shevchik G, Paone S, Dolezal C, Cervenak J. eVisit: a pilot study of a new kind of healthcare delivery. Stud Health Technol Inform 2010; 160: 262–66.
11. Mobile medical applications. Silver Spring, MD: US Food & Drug Administration, Aug10,2018.https://www.fda.gov/MedicalDevices/DigitalHealth/MobileMedicalApplications/default.htm (accessed Feb 20, 2019).

Purpose and Objectives

This is a single-blind, parallel (two-arm), randomized controlled, prospective, non-inferiority trial aiming to assess the efficacy and safety of a proposed artificial-pancreas-like algorithm (AP-A) in MDI therapy in type 2 diabetes.
Hypothesis: Adopting AP-A in guiding pre-prandial insulin titration is non-inferior to the standard clinical approach provided by diabetes specialists from academic center.

Study Population

Sample Size:

We hypothesized that use of the AP-A would not be inferior to physician-guided recommendations in percentage of sensor readings within the target range of 3.9–10 mmol/L. For the sample size calculation, data from the JDRF CGM randomized controlled study and the REPLACE-BG study were used to estimate the properties of the primary endpoint and to define the non-inferiority limit. The sample size was calculated using PASS 15.0 for non-inferiority

analysis, with a margin of 7.5% between arms in a parallel study design with a power of 90%, a one-sided significance level of 5% and SD of 13%. The resulting analysis showed that 106 subjects were needed for the study. The sample size was set at 111-116 participants (we rounded it off to about 120), given an estimated 5–10% dropout rate, with randomization in a 1:1 ratio between the two study arms.

Inclusion Criteria:

Type 2 diabetes subjects receiving intensive insulin therapy;

Age ⩾ 18;

Unsatisfied glucose control, HbA1c ⩾ 8.0%.

Exclusion Criteria:

Type 1 diabetes;

Currently pregnant, planning to be pregnant or unwilling and unable to use contraception during the study (female only);

Mental disorders;

Refusion to wear invasive examination equipment;

Reasons for not able to wear intermittently scanned continuous glucose monitoring (isCGM), e.g., serious allergies, skin diseases, etc.;

Participants who have skin lesions, scarring, redness, infection or edema at the sensor application site that may affect the accuracy of the sensor;

Serious comorbidities, including but not limited to heart disease, cerebrovascular disease, liver and kidney disease, and serious diabetes complications;

Anticipated X-ray, MRI or CT examination during the study;

Usage of drugs that affect blood glucose, such as glucocorticoids, one-month prior enrollment;

Any other reason that investigators assessed to be unsuitable to participate in the study.

Design

This study aims to verify the effecacy and safety of the AP-A in pre-prandial insulin titration for type 2 diabetes on MDI therapy. In the intervention group, the recommended dose of insulin injection was given by the algorithm and implemented after the physician's approval. The control

group was given insulin injection dose directly by the physician. The differences in blood glucose control level between the two groups were compared (determined by TIR and TBR). At the same time, the differences between the actual administered dose and the algorithm recommended dose after the physician's approval were also compared.

Study Procedures

- Type 2 diabetes patients who meet inclusion/exclusion criteria will be offered participation in the study. The research physician would explain the informed consent form and answer relevant questions. Written informed consent form should be signed before the enrollment. Consent will be obtained in a private examination room at the clinic where the patient is receiving care.
- Baseline clinical information will be collected by the research physician including age, sex, diabetes duration, comorbidities, complications, and concomitant medications.
- Height, weight, blood pressure, and waist circumference are measured.
- After an overnight fasting of 10 to 12 hours, fast venous blood (6:00 AM) is collected to measure fasting plasma glucose, creatinine, alanine aminotransferase, aspartate aminotransferase, uric acid, total cholesterol, triglycerides, high-density lipoprotein cholesterol, low-density lipoprotein cholesterol, fasting serum insulin and C-peptide, and HbA1c. Urinary sample is also collected.
- Participants will be randomly assigned to the intervention arm or the control group according to their random number. Participants are masked to the arm assignment, whereas physicians are not masked due to the nature of the intervention.
- Research nurse will place an isCGM sensor for each participant. isCGM wearing methods are as follows:
  Location selection: select the dorsal part of the upper arm (note: avoid scars, moles, pregnancy or obesity marks, lumps, and insulin injection sites). To prevent skin irritation, replace the part when applying the sensor again.
  Cleaning: Clean the application area with alcohol cotton. Allow the applied area to dry before continuing the other steps.

Prepare the sensor applicator: completely tear the packaging film from the sensor assembly package, and unscrew the sensor applicator cover (note: the sensor code must match the sensor assembly package and the sensor applicator).

Align the black mark: Align the black mark on the sensor applicator with the black mark on the sensor assembly package. Press the sensor applicator firmly until it stops.

Lift the sensor applicator from the sensor assembly package.

The sensor applicator is ready to apply the sensor.

Application: Place the sensor applicator on the application part, and press down firmly to apply the sensor (Note: Do not press down before placing the sensor applicator on the prepared application part, to prevent unexpected results or injuries).

Gently remove the sensor applicator from the patient.

Ensure that the sensor is firmly applied. Discard the used sensor applicator and sensor assembly package.

- In the intervention AP-A arm, isCGM data are collected before meal and uploaded into the proposed closed-loop meal-bolus decision algorithm, which could automatically output suggested insulin dosage. The physician would review the suggested dosage and choose either to adopt it, or override the recommended dosage. In the physician arm, insulin dosage adjustment is determined by certified endocrinologists according to current clinical strategies based on isCGM readings.
- The titration period will depend on the actual hospitalization days. We assume that the hospitalization period is 7 days, and the titration period will be 7 days (Figure 5).

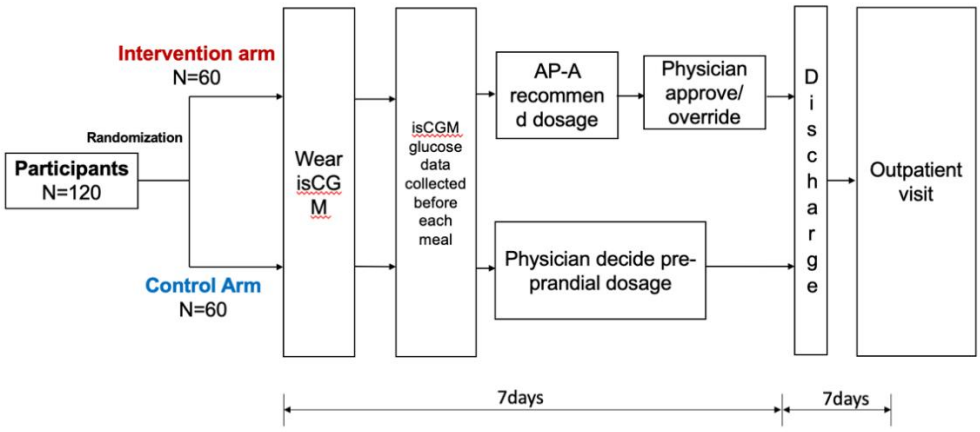


Figure 5. Study procedure

- After the participants are discharged from the hospital, both arms are treated according to physician's instructions.
- Outpatient appointment visit 14 days after enrollment. During the visit, isCGM data will be downloaded and instructions according to the isCGM data will be shared with participants. The participants will then be treated according to routine treatment. 2 ml of blood will be taken for examination of glycosylated albumin. And the study follow-up is over.

Statistical Methods, Data Analysis and Interpretation

Sample Size and Randomization:

We hypothesized that use of the AP-A would not be inferior to physician-guided recommendations in percentage of sensor readings within the target range of 3.9–10 mmol/L. For the sample size calculation, data from the JDRF CGM randomized controlled study and the REPLACE-BG study were used to estimate the properties of the primary endpoint and to define the non-inferiority limit. The sample size was calculated using PASS 15.0 for non-inferiority analysis, with a margin of 7.5% between arms in a parallel study design with power of 90%, a one-sided significance level of 5% and SD of 13%. The resulting analysis showed that 106 subjects were needed for the study. The sample size was set at 111-116 participants (we rounded it off to about 120), given an estimated 5–10% dropout rate, with randomization in a 1:1 ratio between the two study arms.

Eligible patients were randomly assigned (1:1) to AP-A or physician arm. The randomization was stratified by age and HbA_{1c} at screening using a computer-generated random sequence, and had random block size.

Outcomes

Primary endpoints

The primary endpoint was the non-inferiority for percentage time spent with sensor glucose level in 3.9-10.0 mmol/L (70-180 mg/dL) between the two study arms.

The primary safety endpoint was the percentage time spent with sensor glucose level below 3.0 mmol/L (54 mg/dL).

Secondary endpoints

1. Glycated albumin level at the last visit.
2. Proportion of isCGM stored glucose data<3.9mmol/L during treatment adjustment.
3. Proportion of isCGM stored glucose data<2.8mmol/L during treatment adjustment.
4. Proportion of isCGM stored glucose data > 10.0mmol/L during treatment adjustment.
5. Proportion of glucose stored in isCGM > 13.3mmol/L during treatment adjustment.
6. The area of glucose in AGP map under the curve of<10.0mmol/L during treatment adjustment.
7. The area of glucose in AGP map under the curve of<3.9 mmol/L during treatment adjustment.
8. Mean value of isCGM stored glucose data during treatment adjustment (MEAN).
9. Standard deviation (SD) of isCGM stored glucose data during treatment adjustment.
10. Differences between physician approved dose and algorithm recommended dose in intervention group.
11. Number of physician overrides in the intervention group.
12. In the intervention group, if physician overrides occurred, the proportion of glucose within the range of 3.9mmol/L to 10.0mmol/L within 4 hours after insulin injection.

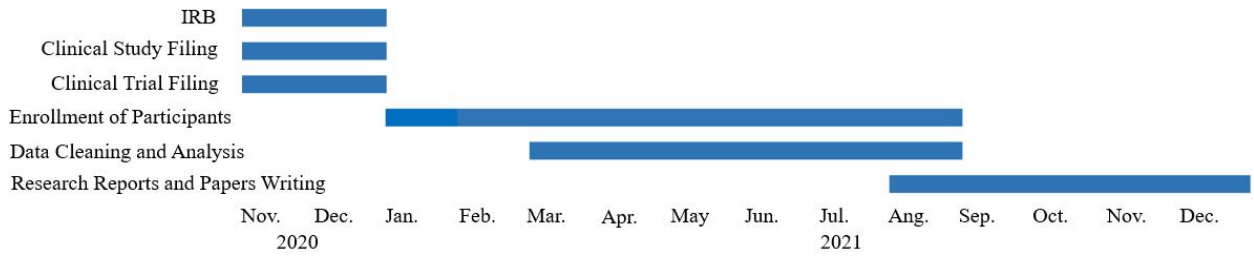
Data Analyses

The analyses will be performed in the intention-to-treat participants. And per-protocol analysis will also be performed. The co-primary end points used a stepdown strategy designed to maintain type I error at 5% or less for multiple comparisons. The AP-A guided arm was considered non-inferior compared to the control arm if co-primary end points were met. A non-inferiority test was conducted for the other endpoints using a randomization test with a significance level of 5%. The non-inferiority was assessed based on the defined margin. For the endpoints related to the hypoglycemia and describable by time in range, the non-inferiority limits were all selected as 2%. The secondary endpoints were then tested in a prespecified order until one of the null hypotheses was nonrejected. The physician's overrides analysis was computed as the percentage of change from the original algorithm recommendations. The statistical analyses

for baseline characteristics of participants were conducted using SPSS, version 27 (IBM Software). Statistical significance was defined as p values less than 0.05 with two-sided testing.

Time schedule of study

The time schedule of study will be organized as follows:



Privacy protection for participants

In order to protect the privacy of the participants, the researchers will use their initials instead of original name during data entry and data analysis, and the data entry does not involve any other content related to their privacy. All data will be saved in the USB flash disk with password.

Summary of changes

1. The protocol has not been changed since the IRB approval (15/Jan/2021).
2. We made one modification of the protocol since it was approved (12/Oct/2020) according to the IRB amendments: adding “Privacy protection for participants” part in the protocol.

北京大学人民医院
PEKING UNIVERSITY PEOPLE'S HOSPITAL

Statistical Analysis Plan: Original

Date: 27 July 2022 (original version 15 Jan 2021)

Study Biostatistician:

Liyuan Tao

Liyuan Tao
Research Center of Clinical Epidemiology, Peking University Third Hospital

Study: RCT assessing efficacy and safety of an artificial-pancreas-like closed-loop control in adults with type 2 diabetes on multiple daily injections therapy

Department of Endocrinology and Metabolism
Peking University People's Hospital
No 11, Xizhimen Nan Street, Xicheng District, Beijing, China, 100044
Tel: +86-10-88324108
Fax: +86-10-88324371

Original Version (from final version of IRB approvement, 15 Jan 2021)

1 Study aim

This single-blind, parallel (two-arm), randomized controlled, prospective, non-inferiority trial aims to determine whether pre-prandial insulin dose adjustments guided by the artificial-pancreas-like algorithm (AP-A) under the supervision of physicians is as effective and safe as those guided by physicians alone.
Hypothesis: Adopting AP-A in guiding pre-prandial insulin titration is non-inferior to the standard clinical approach provided by diabetes specialists from academic center.

2 Ethical approval

This study will be conducted in Peking University People's Hospital, Beijing, China. This study has been registered with Chictr.org (registration number: ChiCTR2200055328), and approved by the institutional review board at Peking University People's Hospital (2020PHB338-01).

3 Sample size

The main study is a randomized controlled trial with noninferiority design. Participants will be assigned 1:1 to either the AP-A arm or the physician arm. We hypothesized that use of the AP-A would not be inferior to physician-guided recommendations in percentage time of sensor readings within the target range of 3.9–10 mmol/L. For the sample size calculation, data from the JDRF CGM randomized controlled study and the REPLACE-BG study were used to estimate the properties of the primary endpoint and to define the non-inferiority limit [1][2]. The sample size was calculated using PASS15.0 for non-inferiority analysis, with a margin of 7.5% between arms in a parallel study design with power of 90%, a one-sided significance level of 5% and SD of 13%. Besides, we assumed that there is no difference between the two arms in achieving the percentage time of sensor readings within the target range of 3.9–10 mmol/L. The resulting analysis showed that 106 subjects were needed for the study. The sample size was set at 111-116 participants (which is rounded off to 120), given an estimated 5–10% dropout rate, with randomization in a 1:1 ratio between the two study arms.

4 Statistical analysis

The study's primary efficacy endpoint is the non-inferiority for percent time spent with sensor glucose level in 3.9-10.0 mmol/L (70-180 mg/dL) between the two study arms. The primary safety endpoint is the percent time spent with sensor glucose level below 3.0 mmol/L (54 mg/dL). For the co-primary endpoint comparisons, the statistical analysis will be performed to evaluate the differences between the two groups during insulin titration period, respectively. Specifically, the mean (95% confidence intervals, CIs) of between-group differences will be calculated. For the

primary efficacy endpoint, non-inferiority will be established if the lower bound of the CI does not less than the noninferiority margin of -7.5%. For the primary safety endpoint, non-inferiority will be established if the upper bound of the CI does not greater than the non-inferiority margin of 2%, which is selected according to previous reports [1][3]. The co-primary end points use a stepdown strategy designed to maintain type I error at 5% or less for multiple comparisons. The secondary endpoints will then be tested in a prespecified order until one of the null hypotheses was non-rejected.

References

1. Tamborlane WV, Beck RW, Bode BW, et al. Continuous glucose monitoring and intensive treatment of type 1 diabetes. N Engl J Med 2008; 359(14):1464-1476.
2. Aleppo G, Ruedy KJ, Riddlesworth TD, et al. REPLACE-BG: a randomized trial comparing continuous glucose monitoring with and without routine blood glucose monitoring in adults with well-controlled type 1 diabetes. Diabetes Care 2017; 40:538–545.
3. Nimri R, Battelino T, Laffel LM, et al. Insulin dose optimization using an automated artificial intelligence-based decision support system in youths with type 1 diabetes. Nat Med 2020;26(9):1380-1384.

北京大学人民医院
PEKING UNIVERSITY PEOPLE'S HOSPITAL

Statistical Analysis Plan: Final

Date: 22 Oct 2022

Study Biostatistician:

Liyuan Tao

Liyuan Tao
Research Center of Clinical Epidemiology, Peking University Third Hospital

Study: RCT assessing efficacy and safety of an artificial-pancreas-like closed-loop control in adults with type 2 diabetes on multiple daily injections therapy

Department of Endocrinology and Metabolism
Peking University People's Hospital
No 11, Xizhimen Nan Street, Xicheng District, Beijing, China, 100044
Tel: +86-10-88324108
Fax: +86-10-88324371

1 Study aim

This single-blind, parallel (two-arm), randomized controlled, prospective, non-inferiority trial aims to determine whether pre-prandial insulin dose adjustments guided by the artificial-pancreas-like algorithm (AP-A) under the supervision of physicians is as effective and safe as those guided by physicians alone.
Hypothesis: Adopting AP-A in guiding pre-prandial insulin titration is non-inferior to the standard clinical approach provided by diabetes specialists from academic center.

2 Methods

2.1 Ethical Approval

This study will be conducted in Peking University People's Hospital, Beijing, China. This study has been registered with Chictr.org (register number: ChiCTR2200055328), and approved by the institutional review board at Peking University People's Hospital (2020PHB338-01).

2.2 Study design

This is a single-blind, parallel (two-arm), randomized controlled, prospective, non-inferiority study. The participants will be randomly divided in 1:1 ratio into AP-A arm and the physician arm. This study aims to assess the efficacy and safety of a proposed AP-A in multiple daily injection (MDI) therapy in type 2 diabetes. In the AP-A arm, the recommended dose of insulin injection was given by the algorithm and implemented after the doctor's approval. The physician arm was given insulin injection dose directly by the physician.

2.3 Randomization and Blinding

After signing the informed consent forms, the participants will be enrolled according to the inclusion and exclusion criteria. The participants who meet the eligibility criteria will be randomized into two arms. Local investigators randomly allocated participants after their enrolment. The randomization was stratified by age and glycated hemoglobin A1c (HbA1c) at screening using a computer-generated random sequence, and had random block size. The participants and statisticians will be masked to the arm assignment, whereas physicians are not masked due to the nature of the intervention.

3 Study outcomes

3.1 Primary endpoints

The primary endpoint is the non-inferiority for percentage time spent with sensor glucose level in 3.9-10.0 mmol/L (70-180 mg/dL) between the two study arms. The primary safety endpoint is the percentage time spent with sensor glucose level below 3.0 mmol/L (54 mg/dL).

3.2 Secondary endpoints

1) Proportion of isCGM stored glucose data < 3.9mmol/L during treatment adjustment.
2) Proportion of isCGM stored glucose data < 2.8mmol/L during treatment adjustment.
3) Proportion of isCGM stored glucose data > 10.0mmol/L during treatment adjustment.
4) Proportion of glucose stored in isCGM > 13.3mmol/L during treatment adjustment.
5) The area of glucose in AGP map under the curve of < 10.0mmol/L during treatment adjustment.
6) The area of glucose in AGP map under the curve of < 3.9 mmol/L during treatment adjustment.
7) Mean value of isCGM stored glucose data during treatment adjustment (MEAN).
8) Standard deviation (SD) of isCGM stored glucose data during treatment adjustment.
9) Differences between physician approved dose and algorithm recommended dose in intervention group.
10) Number of physician overrides in the intervention group.
11) In the intervention group, if physician overrides occurred, the proportion of glucose within the range of 3.9mmol/L to 10.0mmol/L within 4 hours after insulin injection.
12) In the intervention group, if physician overrides occurred, the sensor glucose level at 4 hours after insulin injection.

4 General considerations

4.1 Population

4.1.1 Sample size

The main study is a randomized controlled trial with noninferiority design. Participants will be assigned 1:1 to either the AP-A arm or the physician arm. We hypothesized that use of the AP-A would not be inferior to physician-guided recommendations in percentage of sensor readings within the target range of 3.9–10 mmol/L. For the sample size calculation, data from the JDRF CGM randomized controlled study and the REPLACE-BG study were used to estimate the properties of the primary endpoint and to define the non-inferiority limit [1][2]. The sample size was calculated using PASS15.0 for non-inferiority analysis, with a margin of 7.5% between arms in a parallel study design with power of 90%, a one-sided significance level of 5% and SD of 13%. Besides, we assume that there is no difference between the two arms in achieving the percentage time of sensor readings within the target range of 3.9–10 mmol/L. The equation for calculating the sample size is as follow [3][4]:

$$n = 2\frac{(z_{1-\alpha} + z_{1-\beta})^2 S^2}{(\Delta - \delta)^2}$$

In the equation, n is the sample size of each arm, Δ is the non-inferiority margin of the

primary endpoint, δ is the mean difference at which the power is computed, α is the type I error, β is the type II error, and S^2 is the polled standard deviation of both comparison arms. The resulting analysis showed that 106 subjects were needed for the study. Finally, the study enrolled 119 participants, which met the sample size requirement.

4.1.2 Full analysis set

The major aim of this study is to investigate whether pre-prandial insulin dose adjustments guided by the AP-A under the supervision of physicians is as effective and safe as those guided by physicians alone. The primary endpoint of this study is the percentage of sensor readings within the target range of 3.9–10 mmol/L. Therefore, the full analysis set are defined as all participants who were randomized. The assessment of primary endpoint will be performed based on full analysis set.

4.1.3 Per protocol set

Participants who receive randomization and underwent multiple daily injection therapy, and complete the study according to the study protocol will be included in the per protocol set. The per protocol set is mainly used for sensitivity analysis.

4.2 Covariates analysis

As this is an RCT study, the probability of imbalance in baseline data between the two arms is speculated to be 5% (a minor probability event). Therefore, no multivariate analysis with the adjustment of covariates will be performed to analyze primary endpoint.

4.3 Missing data

Regarding the primary endpoint, the isCGM sensor is wore and collected carefully during MDI therapy or during hospitalization, and thus the probability of missing data associate with primary endpoint remained relatively low. However, if missing data were found, then the percentage of missing data will be reported, the potential patterns of missing data should be examined, and appropriate method should be used for imputation of missing data. The multiple imputation method will be preferred for analyzing the missing data, and the results should be reported in the manuscript as sensitivity analysis.

4.4 Interim analysis

No interim analysis is planned in this study.

5 Statistical analysis

5.1 Data management and general analysis

Electronic dataset system will be used for data collection and management. The data analyses mainly include statistical description and statistical inference. Quantitative data will be described by central tendency and dispersion tendency. The normally distributed data of central tendency and dispersion tendency will be described as means and standard division, respectively. The non-normally distributed data of central tendency and dispersion tendency will be described as median and quartiles. The qualitative data will be described as frequency and percentage. Statistical inference, independent *t* test or non-parameter test will be used to compare the quantitative data between the two arms, while chi-square test or Fisher's exact test will be used for comparing the qualitative data between the two arms.

5.2 Analysis of primary endpoints

For the co-primary endpoint comparisons, the statistical analysis will be performed to evaluate the differences between the two groups for the two primary endpoints during MDI therapy. Specifically, the co-primary endpoints used a stepdown strategy designed to maintain type I error at 5% or less for multiple comparisons. The primary endpoint followed no Gaussian distribution, will be presented as median (interquartile range). The mean (95% confidence intervals, CIs) of between-group differences of the median will be calculated by bootstrap method (1000 replications). For the primary efficacy endpoint, non-inferiority will be established if the lower bound of the CI does not less than the noninferiority margin of -7.5%. For the primary safety endpoint, non-inferiority will be established if the upper bound of the CI does not greater than the non-inferiority margin of 2%, which is selected according to previous reports [1][5]. The per-protocol (PP) sets will be also used for sensitivity analysis.

5.3 Analysis of other endpoints

For secondary endpoints, continuous data will be presented as means (SDs) or median (IQRs), as appropriate. The qualitative data will be described as frequency and percentage. Specifically, the isCGM derived secondary endpoints from 1) to 8) in Section 3.2 followed no Gaussian distribution, will be presented as median (interquartile range). The mean (95% CIs) of between-group differences of the median will be calculated by bootstrap method (1000 replications). For the endpoints related to the hypoglycemia and describable by time in range, the non-inferiority limits were all selected as 2% [5]. The secondary endpoints will then be tested in a prespecified order until one of the null hypotheses was nonrejected. Non-inferiority is met when the upper confidence limit is greater than the non-inferiority margin or when the lower confidence limit is less than the non-inferiority margin.

For the secondary endpoint described in 9), the differences between physician approved dose and algorithm recommended dose in intervention group will be represented using Bland-Altman plots, where mean value of the differences, the 95% limits of agreement together with the proportion of the differences falling in the limits will also be calculated. For the secondary endpoint described in 10), the total number of AP-A recommendations made for participants and physician overrides in the

intervention group will be collected, respectively. Besides, the reasons for the physician overrides will also be provided. For the secondary endpoints described in 11) and 12), we will divide the total dosages suggested by AP-A into three parts: in consistent with physician approvement, greater than actual administered dosage, and less than actual administered dosage. For each part and physician arm, the percent time spent with sensor glucose in 3.9-10.0 mmol/L within 4 hours after insulin injection, and sensor glucose at 4 hours after insulin injection will be calculated. If the outcomes do not follow Gaussian distribution, they will be presented as median (interquartile range). Besides, the between-part differences will also be analyzed using Mann-Whitney U test. Specifically, the analyzed differences are presented in the following items:

- When the recommended insulin dosages are greater than actual administered ones, the percent time spent with sensor glucose in 3.9-10.0 mmol/L within 4 hours after insulin injection will be compared with the circumstances consistent with physician decisions in the AP-A arm and that in the physician arm using Mann-Whitney U test, respectively.
- When the recommended insulin dosages are greater than actual administered ones, the sensor glucose at 4 hours after insulin injection will be compared with the circumstances consistent with physician decisions in the AP-A arm and that in the physician arm using Mann-Whitney U test, respectively.
- When the recommended insulin dosages are less than actual administered ones, the percent time spent with sensor glucose in 3.9-10.0 mmol/L within 4 hours after insulin injection will be compared with the circumstances consistent with physician decisions in the AP-A arm and that in the physician arm using Mann-Whitney U test, respectively.
- When the recommended insulin dosages are less than actual administered ones, the sensor glucose at 4 hours after insulin injection will be compared with the circumstances consistent with physician decisions in the AP-A arm and that in the physician arm using Mann-Whitney U test, respectively.

5.4 Graphical analysis

To clearly present the between-arm differences of the endpoints, a forest plot will also be provided for the isCGM derived endpoints. Besides, a comparison plot of isCGM glucose pattern between the two arms during the study period will also be presented, where the hypoglycemia events (< 3.0 mmol/L (54 mg/dL)), median, areas enclosed by 25%-75% and 5%-90% quantiles of sensor glucose level of each arm will be plotted. Finally, to compare the insulin dosage recommended by AP-A with the actual executed dosage, a Bland-Altman plot of the agreement between proposed and actual administered insulin dosage will be provided, together with a plot of the differences between proposed and actual administered insulin dosage during the titration period for each participant, where the differences between proposed and actual administered insulin dosage will be represented as dots, and the sizes of the dots will be proportional to their frequency. Besides, the median line (smoothed) of differences during the titration period will also be provided. Finally, the zone enclosed by the

10% and 90% quartile lines (smoothed) of the differences during the titration period will also be presented, and the number of the differences in each titration day will be identified in the plot. The pseudo-codes for the aforementioned plots are provided as follows:

Algorithm 1 Pseudo-code of the forest plot using R

Require: Data of the isCGM derived endpoints for each arm, data of the between-arm differences (mean (95% CIs)) of the isCGM derived endpoints
Ensure: A forest plot for the isCGM derived endpoints
Read data as fp ← read.csv("data address") in R
Determine the style of forest plot Data ← fpShapesGp(lines, box) in R
Output forest plot ← forestplot(fp, options, style) in R
return

Algorithm 2 Pseudo-code of the comparison plot of isCGM glucose pattern between the two arms using Matlab

Require: isCGM data files for the two arms
Ensure: Comparison plot of isCGM glucose pattern between the two arms
for each file in the AP-A arm do
    Get isCGM data
    Get isCGM data sampling time
    for $i = 1$ to the length of isCGM data
        if ith isCGM data < 54 mg/dL
            $n_a \leftarrow n_a \cup i$th isCGM data sampling time
        end
        Determine the location of ith isCGM data in the day (24h) according to its sampling time
        Fill the *NaN* value in the location of vacant isCGM data in the day
        Get daily isCGM data in chronological order stored in the rows of a matrix φ
    end
    $\varphi_a \leftarrow \varphi_a \cup \varphi$
end
for each file in the physician arm do
    Get isCGM data
    Get isCGM data sampling time
    for $i = 1$ to the length of isCGM data
        if ith isCGM data < 54 mg/dL
            $n_b \leftarrow n_b \cup i$th isCGM data sampling time
        end
        Determine the location of ith isCGM data in the day (24h) according to its sampling time
        Fill *NaN* value in the location of vacant isCGM data in the day
        Get daily isCGM data in chronological order stored in the rows of a matrix φ
    end
    $\varphi_b \leftarrow \varphi_b \cup \varphi$
end
Compute median, 5%, 25%, 75% and 90% quartile values of sensor glucose level in the day for AP-A arm using φ_a
Compute median, 5%, 25%, 75% and 90% quartile values of sensor glucose

level in the day for physician arm using φ_b
29: Plot areas enclosed by 25%-75% and 5%-90% quartile lines of sensor glucose level in the day for AP-A arm, respectively
30: Plot a median line of sensor glucose level in the day for AP-A arm
31: Plot hypoglycemia events for AP-A arm using n_a
32: Plot areas enclosed by 25%-75% and 5%-90% quartile lines of sensor glucose level in the day for physician arm, respectively
33: Plot a median lines of sensor glucose level in the day for physician arm
34: Plot hypoglycemia events for physician arm using n_b

Algorithm 3 Pseudo-code of Bland-Altman plot using Matlab

Require: Data of the differences between proposed and actual administered insulin dosage
Ensure: A Bland-Altman plot of the agreement between proposed and actual administered insulin dosage
1: Compute the mean value m of the differences
2: Compute the standard deviation value SD of the differences
3: Compute upper value of 95% limits of agreement using m+1.96SD
4: Compute lower value of 95% limits of agreement using $m - 1.96SD$
5: for each difference do
6: if difference < m+1.96SD && difference > $m - 1.96SD$
7: $n \leftarrow n + 1$
8: end
9: end
10: Compute the proportion of the differences falling in the limits using n
11: Plot the differences as dots
12: Plot a mean line of the differences
13: Plot lines of 95% limits of agreement

Algorithm 4 Pseudo-code of the plot of differences between proposed and actual administered insulin dosage during the titration period for each participant using Matlab

Require: Data of the differences between proposed and actual administered insulin dosage during the titration period for each participant
Ensure: A plot of differences during the titration period for each participant
1: Determine the longest titration period among all the participants
2: Generate a $n \times N$ matrix φ that stores all the differences; n is the number of the participants and N is the longest titration period
3: for each participant do
4: Store the differences of each participant in the row of φ
5: Fill the value of *NaN* in the location of vacant differences
6: end
7: for $i = 1$ to N do
8: Count the number of the kind of difference and the number of each kind of differences in $\varphi(:, i)$
9: Plot each kind of differences at abscissa i as a dot whose size is proportional to its number
Count the number of the total differences in each titration day
10: end
11: Compute median, 10%, and 90% quartile values of the differences during the titration period using φ

12: Plot a median line (smoothed) of the differences
13: Plot areas enclosed by 10%-90% quartile lines (smoothed) of the differences
14: Identify the number of the differences in each titration day

5.5 Software and significant level of the statistical analyses

The statistical analyses for baseline characteristics of participants will be conducted using SPSS, version 27 (IBM Software). Other statistical analyses will be performed using MATLAB 2017b (Mathworks, Inc., Natick, MA) for Windows and R, version 4.2.1 (R Foundation for Statistical Computing). Statistical significance is set as $P<0.05$ with two-sided testing. All plots will be generated by the MATLAB 2017b (Mathworks, Inc., Natick, MA), except for the forest plot, which will be provided using R, version 4.2.1.

References

1. Tamborlane WV, Beck RW, Bode BW, et al. Continuous glucose monitoring and intensive treatment of type 1 diabetes. N Engl J Med 2008; 359(14):1464-1476.
2. Aleppo G, Ruedy KJ, Riddlesworth TD, et al. REPLACE-BG: a randomized trial comparing continuous glucose monitoring with and without routine blood glucose monitoring in adults with well-controlled type 1 diabetes. Diabetes Care 2017; 40:538–545.
3. Chow SC, Shao J, Wang H. Sample Size Calculations in Clinical Research (3rd ed.). Chapman and Hall/CRC. 2017.
4. Julious Steven A. Tutorial in Biostatistics. Sample sizes for clinical trials with Normal data. Statistics in Medicine. 2004; 23:1921-1986.
5. Nimri R, Battelino T, Laffel LM, et al. Insulin dose optimization using an automated artificial intelligence-based decision support system in youths with type 1 diabetes. Nat Med 2020;26(9):1380-1384.

Summary of Changes for Statistical Analysis Plan

[1] Amendments at secondary endpoints after the study

One of the secondary endpoints, glycated albumin (GA) level at the last visit was not included in the of the final analysis, since GA was not obtained from more than half of the participants due to frequent partial lock-down because of COVID-19 during the study period in Beijing.
Meanwhile, to further explore the safety of the algorithm, an additional secondary endpoint was included in the final statistical analysis plan. In the intervention group, if physician overrides occurred, the sensor glucose level at 4 hours after insulin injection.

[2] Amendments at the analysis of physician overrides

To further explore the performance of the algorithm, the dosages suggested by AP-A will be divided into three categories: in consistent with physician approvement, greater than actual administered dosage, and less than actual administered dosage. The percent time spent with sensor glucose in 3.9-10.0 mmol/L within 4 hours after insulin injection, and sensor glucose at 4 hours after insulin injection will be calculated in each category and the physician arm. If the outcomes do not follow Gaussian distribution, they will be presented as median (interquartile range) and tested by Mann-Whitney U test.

[3] Graphical analysis

To clearly present the between-arm differences of the endpoints, a forest plot will be provided for the isCGM derived endpoints. Besides, a comparison plot of isCGM glucose pattern between the two arms during the study period will also be presented, in which the hypoglycemia events (< 3.0 mmol/L (54 mg/dL)), median, 25%-75% and 5%-90% sensor glucose level of each arm will be plotted. Finally, to compare the insulin dosage recommended by AP-A with the actual executed dosage, a Bland-Altman plot of the agreement between proposed and actual administered insulin dosage will be presented, together with a plot of the differences between proposed and actual administered insulin dosage during the titration period for each participant.