

## Principal Investigator

**First Name:** Qi  
**Last Name:** Pan  
**Degree:** MD, PhD  
**Primary Affiliation:** Beijing Hospital  
**E-mail:** [panqi621@126.com](mailto:panqi621@126.com)  
**State or Province:** Beijing  
**Country:** China

## General Information

### Key Personnel (other than PI):

**First Name:** Jia  
**Last name:** Zhang  
**Degree:** MD  
**Primary Affiliation:** Beijing Hospital  
**SCOPUS ID:**  
**Requires Data Access?** Yes

**First Name:** Weihao  
**Last name:** Wang  
**Degree:** MD  
**Primary Affiliation:** Beijing Hospital  
**SCOPUS ID:**  
**Requires Data Access?** Yes

**Are external grants or funds being used to support this research?:** External grants or funds are being used to support this research.

**Project Funding Source:** Beijing Municipal Science & Technology Commission (Z221100007422007), Capital's Funds for Health Improvement and Research (2022-1-4051), Noncommunicable Chronic Diseases-National Science and Technology Major Project (2023ZD0508800), Bethune Charitable Foundation (Z04J2024E120)

**How did you learn about the YODA Project?:** Internet Search

## Conflict of Interest

[https://yoda.yale.edu/wp-content/uploads/2025/02/JiaZhang\\_disclosure-2.pdf](https://yoda.yale.edu/wp-content/uploads/2025/02/JiaZhang_disclosure-2.pdf)  
[https://yoda.yale.edu/wp-content/uploads/2025/02/QiPan\\_disclosure-2.pdf](https://yoda.yale.edu/wp-content/uploads/2025/02/QiPan_disclosure-2.pdf)  
[https://yoda.yale.edu/wp-content/uploads/2025/02/WeihaoWang\\_disclosure-2.pdf](https://yoda.yale.edu/wp-content/uploads/2025/02/WeihaoWang_disclosure-2.pdf)

## Certification

**Certification:** All information is complete; I (PI) am responsible for the research; data will not be used to support litigious/commercial aims.

**Data Use Agreement Training:** As the Principal Investigator of this study, I certify that I have completed the YODA Project Data Use Agreement Training

1. [NCT00642278 - 28431754DIA2001 - A Randomized, Double-Blind, Placebo-Controlled, Double-Dummy, Parallel Group, Multicenter, Dose-Ranging Study in Subjects With Type 2 Diabetes](https://clinicaltrials.gov/ct2/show/NCT00642278)

[Mellitus to Evaluate the Efficacy, Safety, and Tolerability of Orally Administered SGLT2 Inhibitor JNJ-28431754 With Sitagliptin as a Reference Arm](#)

2. [NCT01106625 - 28431754DIA3002 - A Randomized, Double-Blind, Placebo-Controlled, 3-Arm, Parallel-Group, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Control on Metformin and Pioglitazone Therapy](#)

3. [NCT01064414 - 28431754DIA3004 - A Randomized, Double-Blind, Placebo-Controlled, 3-Arm, Parallel-Group, 26-Week, Multicenter Study With a 26-Week Extension, to Evaluate the Efficacy, Safety and Tolerability of Canagliflozin in the Treatment of Subjects With Type 2 Diabetes Mellitus Who Have Moderate Renal Impairment](#)

4. [NCT01081834 - 28431754DIA3005 - A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin as Monotherapy in the Treatment of Subjects With Type 2 Diabetes Mellitus Inadequately Controlled With Diet and Exercise](#)

5. [NCT01106677 - 28431754DIA3006 - A Randomized, Double-Blind, Placebo and Active-Controlled, 4-Arm, Parallel Group, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Control on Metformin Monotherapy](#)

6. [NCT00968812 - 28431754DIA3009 - A Randomized, Double-Blind, 3-Arm Parallel-Group, 2-Year \(104-Week\), Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of JNJ-28431754 Compared With Glimepiride in the Treatment of Subjects With Type 2 Diabetes Mellitus Not Optimally Controlled on Metformin Monotherapy](#)

7. [NCT01106651 - 28431754DIA3010 - A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin Compared With Placebo in the Treatment of Older Subjects With Type 2 Diabetes Mellitus Inadequately Controlled on Glucose Lowering Therapy](#)

8. [NCT01106690 - 28431754DIA3012 - A Randomized, Double-Blind, Placebo-Controlled, 3-Arm, Parallel-Group, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Control on Metformin and Sulphonylurea Therapy](#)

9. [NCT01137812 - 28431754DIA3015 - A Randomized, Double-Blind, Active-Controlled, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin Versus Sitagliptin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Control on Metformin and Sulphonylurea Therapy](#)

10. [NCT01809327 - 28431754DIA3011 - A Randomized, Double-Blind, 5-Arm, Parallel-Group, 26-Week, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin in Combination With Metformin as Initial Combination Therapy in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Control With Diet and Exercise](#)

11. [NCT01381900 - 28431754DIA3014 - A Randomized, Double-Blind, Placebo-Controlled, Parallel Group, 18-Week Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Control on Metformin Alone or in Combination With a Sulphonylurea](#)

12. [NCT01340664 - 28431754DIA2003 - A Randomized, Double-Blind, Placebo-Controlled, 3-Arm, Parallel-Group, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Control on Metformin](#)

13. [NCT02025907 - 28431754DIA4004 - A Randomized, Double-blind, Placebo Controlled, 2-arm, Parallel-group, 26-week, Multicenter Study to Evaluate the Efficacy, Safety, and Tolerability of Canagliflozin in the Treatment of Subjects With Type 2 Diabetes Mellitus With Inadequate Glycemic Control on Metformin and Sitagliptin Therapy](#)

14. [NCT01032629 - 28431754DIA3008 - A Randomized, Multicenter, Double-Blind, Parallel, Placebo-Controlled Study of the Effects of JNJ-28431754 on Cardiovascular Outcomes in Adult Subjects With Type 2 Diabetes Mellitus](#)

15. [NCT01989754 - 28431754DIA4003 - A Randomized, Multicenter, Double-Blind, Parallel, Placebo-Controlled Study of the Effects of Canagliflozin on Renal Endpoints in Adult Subjects With Type 2 Diabetes Mellitus](#)

16. [NCT00650806 - 28431754OBE2001 - A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group, Dose-Ranging Study to Investigate the Safety and Efficacy of JNJ-28431754 in](#)

[Nondiabetic Overweight and Obese Subjects](#)

17. [NCT02243202 - 28431754OBE2002 - A Randomized, Double-Blind, Placebo-Controlled, Parallel-Group Study to Investigate the Safety and Efficacy of the Co-administration of Canagliflozin 300 mg and Phentermine 15 mg Compared With Placebo for the Treatment of Non-diabetic Overweight and Obese Subjects](#)
18. [NCT02065791 - 28431754DNE3001 - A Randomized, Double-blind, Event-driven, Placebo-controlled, Multicenter Study of the Effects of Canagliflozin on Renal and Cardiovascular Outcomes in Subjects With Type 2 Diabetes Mellitus and Diabetic Nephropathy](#)
19. [NCT03267576 - 28431754DIA4026 - Canagliflozin Continuous Glucose Monitoring \(CANA CGM\) Trial: A Pilot Randomized, Double-Blind, Controlled, Crossover Study on the Effects of the SGLT-2 Inhibitor Canagliflozin \(vs. the DPP-4 Inhibitor Sitagliptin\) on Glucose Variability in Mexican Patients With Type 2 Diabetes Mellitus Inadequately Controlled on Metformin](#)
20. [NCT02139943 - 28431754DIA2004 - A Randomized Phase 2, Double-blind, Placebo-controlled, Treat-to-Target, Parallel-group, 3-arm, Multicenter Study to Assess the Efficacy and Safety of Canagliflozin as Add-on Therapy to Insulin in the Treatment of Subjects With Type 1 Diabetes Mellitus](#)

**What type of data are you looking for?:** Individual Participant-Level Data, which includes Full CSR and all supporting documentation

## Research Proposal

### Project Title

Tree-like representation of heterogeneity in response to canagliflozin in diabetes.

### Narrative Summary:

This study explores phenotypic heterogeneity in diabetes and its relationship with canagliflozin treatment outcomes, using a tree-like structure generated by the DDTTree algorithm. By mapping patients to the tree-like structure, we aim to identify individuals that benefit most from canagliflozin in terms of efficacy outcomes and cardiorenal outcomes. Additionally, we will compare the DDTTree algorithm's performance with traditional k-means clustering to evaluate the effectiveness of each method in capturing phenotypic variation. The goal is to enhance precision medicine in diabetes by improving patient stratification, ultimately guiding clinical decisions and optimizing treatment efficacy.

### Scientific Abstract:

**Background:** Phenotypic heterogeneity in diabetes, particularly in response to treatment, can be visualized using a tree-like structure through the DDTTree algorithm. This method offers potential to stratify individuals based on phenotypic characteristics and assess the differential efficacy of diabetes treatments. However, the differences in response to canagliflozin, a novel treatment, across the phenotypic tree remain underexplored.

**Objective:** The study aims to investigate whether continuous stratification by the phenotypic tree structure can identify individuals who derive greater benefit from canagliflozin treatment, thereby guiding clinical decision-making.

**Study Design:** A post hoc analysis of data from a canagliflozin clinical trial will be conducted.

**Participants:** The study included individuals from the clinical trial who were treated with canagliflozin and had baseline and follow-up data available for analysis.

**Primary and Secondary Outcome Measures:** The main outcome is the change in glucose and insulin resistance at the endpoint. Cardiovascular outcomes (Major Adverse Cardiovascular Events, MACE) and renal outcomes (progression of albuminuria and changes in urinary albumin/creatinine ratio and eGFR) will be evaluated.

**Statistical Analysis:** The relationship between tree coordinates and outcomes would be assessed using Cox proportional hazards, logistic regression, and linear regression. Additionally, multinomial regression will be used to predict hard clusters on the tree. Stata and R studio will be used as analytical tools.

### **Brief Project Background and Statement of Project Significance:**

Phenotypic heterogeneity and its association with the risk of diabetes outcomes in diabetes can be visualized using a Tree-like structure through the DDTTree algorithm.<sup>1,2</sup> Further investigation is warranted to explore the differences in response to these novel agents, such as canagliflozin, in terms of efficacy and diabetes-related outcomes among individuals treated with them on tree-based structure.

### **Specific Aims of the Project:**

This study aimed to investigate whether continuous stratification of individuals by the phenotypic tree can distinguish those deriving greater benefit from treatment with canagliflozin, and therefore guide clinical decisions, by a post hoc analysis of canagliflozin clinical trial data.

### **Study Design:**

Individual trial analysis

### **What is the purpose of the analysis being proposed? Please select all that apply.**

Other: A Stratification Study for Type 2 Diabetes

## **Research Methods**

### **Data Source and Inclusion/Exclusion Criteria to be used to define the patient sample for your study:**

T2D population with complete data on HbA1c, BMI, high-density lipoprotein cholesterol (HDL-C), triglycerides, total cholesterol (TC), alanine aminotransferase (ALT), creatinine, systolic and diastolic blood pressure (SBP and DBP) at baseline were included. We will perform the analyses on the Rstudio.

### **Primary and Secondary Outcome Measure(s) and how they will be categorized/defined for your study:**

Main Outcome Measure and how it will be categorized/defined for your study:

#### 1. Efficacy outcomes:

Main outcome: the decline in A1c at the endpoint of treatment after baseline A1c was adjusted.

Secondary outcomes include: (1) Hypoglycemic episodes across the trial period. (2) glucose lowering durability: difference between A1c at the endpoint of treatment and the end of treatment (EOT). (3) change in beta cell function and insulin resistance as assessed by HOMA2B or HOMA2IR from baseline at EOT.

#### 2. cardiorenal outcomes:

CV outcome: Major Adverse Cardiovascular Events (MACE) Composite of Cardiovascular (CV) Death, Non-Fatal Myocardial Infarction (MI), and Non-Fatal Stroke in the four cluster groups. Renal outcome: Progression of Albuminuria, either from non-albuminuria to albuminuria or from microalbuminuria to macroalbuminuria.

Secondary outcomes: (1) Change from baseline in urinary albumin/creatinine ratio (ACR), (2) change from baseline in estimated glomerular filtration Rate (eGFR) at EOT.

### **Main Predictor/Independent Variable and how it will be categorized/defined for your study:**

phenotypic variables: Age, Sex, HbA1c, BMI, high-density lipoprotein cholesterol (HDL-C), triglycerides, total cholesterol (TC), alanine aminotransferase (ALT), creatinine, systolic and diastolic blood pressure (SBP and DBP).

Disease history (categorized as Yes/No for each) hypertension, heart failure, hyperlipidemia, cardiovascular disease, microvascular disease (including chronic kidney disease, diabetic retinopathy, neuropathy)

Medication history (categorized as Yes/No for each) insulin, other anti-diabetic drugs, antihypertensive, anti-lipidemia, other cardiovascular drugs  
follow-up time

**Other Variables of Interest that will be used in your analysis and how they will be categorized/defined for your study:**

None.

**Statistical Analysis Plan:**

We will present the clinical characteristics of the included population and use the "map" function to map individuals to the reference tree model. Additionally, we will assess the relationship between tree coordinates and the risk of outcomes through Cox proportional hazards models, logistic regression, and linear regression. To compare the performance of the k-means method 3 and the tree-like structure generated by the DDRTree algorithm, we will apply the k-means clustering method and fit multinomial regression models.

**Software Used:**

RStudio

**Project Timeline:**

Month 1 to 2: Database retrieval data organization.

Month 2-4: data analysis and making conclusions.

Month 4-6: article writing and submitting. Report results to the YODA project at the same time that the manuscript was submitted.

Month 6-12: Making corrections and possible reamendment for reviewer's comments.

**Dissemination Plan:**

We plan to submit a paper naming 'Tree-like representation of heterogeneity in response to canagliflozin in diabetes'. We aim to submit to Lancet diabetes & endocrinology, Diabetes care or JAMA network open. Also, we will submit abstract to the ADA, EASD and IDF and present our results in these conferences

**Bibliography:**

1. Nair ATN, Wesolowska-Andersen A, Brorsson C, et al. Heterogeneity in phenotype, disease progression and drug response in type 2 diabetes. *Nature medicine* 2022; **28**(5): 982-8.
2. Schön M, Prystupa K, Mori T, et al. Analysis of type 2 diabetes heterogeneity with a tree-like representation: insights from the prospective German Diabetes Study and the LURIC cohort. *The lancet Diabetes & endocrinology* 2024; **12**(2): 119-31.
3. Ahlqvist E, Storm P, Käräjämäki A, et al. Novel subgroups of adult-onset diabetes and their association with outcomes: a data-driven cluster analysis of six variables. *The lancet Diabetes & endocrinology* 2018; **6**(5): 361-9.