Patient preferences for second-line Type 2 diabetes medication treatment outcomes

NCT05214573

22July2021

IRB Minimal Risk Protocol Template

Note: If this study establishes a human specimen repository (biobank) for research purposes, do not use this template. Use the Mayo Clinic Human Specimen Repository Protocol Template found on the IRB home page under Forms and Procedures at http://intranet.mayo.edu/charlie/irb/

First-time Use: Use this template to describe your study for a new IRB submission.

1. Complete the questions that apply to your study.
2. Save an electronic copy of this protocol for future revisions.
3. When completing your IRBe application, you will be asked to upload this document to the protocol section.

Modification: To modify this document after your study has been approved:

1. Open your study in IRBe. Click on the study 'Documents' tab and select the most recent version of the protocol. Save it to your files.
2. Open the saved document and activate "Track Changes".
3. Revise the protocol template to reflect the modification points , save the template to your files
4. Create an IRBe Modification for the study and upload the revised protocol template.

General Study Information

Principal Investigator: Rozalina G. McCoy, M.D.

Study Title: Patient preferences for second-line Type 2 diabetes medication treatment outcomes

Protocol version number and date: Version 1; July 22, 2021

Research Question and Aims

Hypothesis: Patients with type 2 diabetes will have different preferences and priorities for what outcomes of diabetes management are most important to them.

Aims, purpose, or objectives: This study is a component of a broader study examining the comparative effectiveness and safety of second-line glucose-lowering medications in the management of type 2 diabetes. All other components of this study (other than this one) use de-identified administrative claims data from OptumLabs Data Warehouse (OLDW) and as such are exempt from IRB review.

For this human subjects study, we will elicit patient preferences toward various treatment outcomes using a participatory ranking exercise and use these rankings to generate individually weighted composite outcomes. These composite outcomes (without any linked participant information) will be used for analyses within OLDW that will estimate patient-centered treatment effects of four different classes of second-line medications for Type 2 diabetes with respect to those composite outcomes specifically. The four drug classes that will be examined

are the GLP-1 receptor agonists (GLP-1RA), SGLT2 inhibitors (SGLT2i), DPP-4 inhibitors (DPP-4i), and sulfonylureas (SU).

Background (*Include relevant experience, gaps in current knowledge, preliminary data, etc.*):

Type 2 diabetes (T2D) is one of the most common and serious chronic health conditions in the U.S.[1] Optimal patient-centered T2D care is predicated on treating each patient with medications that are likely to yield the most benefit and risk the least harm, weighing the best available evidence against each patient's preferences and situation. While metformin is consistently recommended as the 1st-line drug in the management of T2D, there is less clarity regarding 2nd-line therapy as new drug classes were approved and introduced into practice,[2-5] resulting in decisional dilemmas for patients and clinicians, and potential missed opportunities to select the drug that best optimizes the patient's health.

Cardiovascular disease (CVD), heart failure (HF), and chronic kidney disease (CKD) are leading causes of morbidity, mortality, and costs associated with T2D.[6-9] Randomized controlled trials (RCTs) designed to evaluate the CV safety of T2D drugs revealed additional therapeutic benefits of glucagon-like peptide-1 receptor agonists (GLP-1RA) and sodium-glucose cotransporter 2 inhibitors (SGLT2i) with respect to CVD, HF, and CKD outcomes among patients at high risk for these events. However, it is unknown how each patient's preferences and values impact the comparative effectiveness of these drugs.

As a result, patients and clinicians face a decisional dilemma at the heart of T2D management: *Which drug class would be best for the specific set of outcomes desired by an individual patient with T2D?* Addressing these knowledge gaps is critical for patients and clinicians to engage in shared decision-making; for professional societies to develop guidelines that reflect scientific evidence and patient priorities; and for health systems and payers to make formulary decisions.

Study Design and Methods

Methods: *Describe in lay terms, completely detailing the research activities that will be conducted by Mayo Clinic staff under this protocol.*

We will ask a purposive sample of patients with T2D (N=25) to engage in a participatory ranking exercise in which they will rank the outcomes of T2D therapy are most important to them personally and then contextualize these decisions. Participatory ranking exercises are widely used in studies of preference elicitation to determine the relative importance of treatment attributes and outcomes.[10] We will develop an inventory of potential outcomes and adverse events associated with the four drug classes of interest (e.g., hospitalization, severe hypoglycemia). As part of an individual interview with a member of the research team, we will ask patient participants to rank these pre-specified outcomes in order of their relative personal importance using a secure online survey platform (Qualtrics). The researcher will ask patients to contextualize their ranking decisions. Next, participants will be asked to identify additional outcomes that are missing from this prespecified list of clinical outcomes, but which may be important to patients with T2D (e.g., medication cost, frequency of administration). Finally, participants will be asked to incorporate these additional, patient-generated outcomes into their ranking decisions and explain their rationale to the researcher.

We will seek to minimize cognitive burden on participants through iterative feedback and pilot-testing with patient-investigators, using lay language and avoiding medical jargon, identifying an appropriate number of pre-specified outcomes for ranking, and ensuring usability of Qualtrics, a secure online ranking platform. The ranking exercise will be conducted using telephone or videoconferencing software, or in person, depending on patient preference. The ranking exercise will be audio-recorded to ensure an accurate record of participant responses to questions about their ranking choices.

Subsequent to the participatory ranking exercise, all participants will be interviewed using a semi-structured interview guide. The interview guide will ask participants about their previous experiences with diabetes medications, the factors that are most important to them in general when making health care treatment decisions, the types of discussions about treatment options that occur during clinical visits, and any differences they have observed in outcomes that are important to health care providers versus those that are most important to patients.

Patient participants (N=25) will be recruited from primary care and endocrinology practices in Mayo Clinic and Mayo Clinic Health System in Minnesota and Wisconsin. Inclusion criteria are: ≥21 years old, T2D, use of ≥1 study drug (GLP-1RA, SGLT2i, DPP-4i, SU). Exclusion criteria are: insulin use, cognitive impairment, terminal/advanced illness, non-English speaking, residency in a long-term care setting.

A study coordinator at Mayo Clinic will recruit patients using Epic to identify patients who meet eligibility and sampling criteria. The study coordinator will call eligible participants to describe the study, obtain verbal informed consent, obtain HIPAA authorization, schedule the ranking exercise session, and obtain mailing information for the remuneration, which will be mailed after completion of the ranking exercise. Sampling will be guided by the principle of response saturation. We will use stratified purposeful sampling to intentionally select participants representing predefined demographic (age, sex, race/ethnicity, insurance type) and clinical (HbA_{1c} level, T2D drugs used) characteristics.

Study participation will be voluntary, and participants will be compensated with a $25 gift card for taking part in this study.

Subject Information

Target accrual is the proposed total number of subjects to be included in this study at Mayo Clinic. A "Subject" may include medical records, images, or specimens generated at Mayo Clinic and/or received from external sources.

Target accrual: 25

Subject population (children, adults, groups): Adults with Type 2 diabetes receiving care from primary care and endocrinology practices in Mayo Clinic and Mayo Clinic Health System in Minnesota and Wisconsin.

Inclusion Criteria: ≥21 years old, T2D diagnosis, use of ≥1 study drug (GLP-1RA, SGLT2i, DPP-4i, SU)

Exclusion Criteria: Insulin use, cognitive impairment, terminal/advanced illness, non-English speaking, residency in a long-term care setting.

MAYO CLINIC

Biospecimens

Collection of blood samples. When multiple groups are involved copy and paste the appropriate section below for example repeat section b when drawing blood from children and adults with cancer.

a. **From healthy, non-pregnant, adult subjects who weigh at least 110 pounds**. For a minimal risk application, the amount of blood drawn from these subjects may not exceed 550ml in an 8 week period and collection may not occur more frequently than 2 times per week.
   Volume per blood draw: _____ml
   Frequency of blood draw (e.g. single draw, time(s) per week, per year, etc.) _____

b. **From other adults and children considering age, weight, and health of subject.** For a minimal risk application, the amount of blood drawn from these subjects may not exceed the lesser of 50 ml or 3 ml per kg in an 8 week period, and collection may not occur more frequently than 2 times per week.
   Volume per blood draw: _____ml
   Frequency of blood draw (e.g. single draw, time(s) per week, per year, etc.) _____

Prospective collection of biological specimens other than blood: _____

Review of medical records, images, specimens

Check all that apply (data includes medical records, images, specimens).

☐ Only data that exists before the IRB submission date will be collected.

Date Range for Specimens and/or Review of Medical Records:
Examples: *01/01/1999 through 12/31/2015*, or all records through *mm/dd/yyyy*.

Note: The Date Range must include the period for collection of baseline data, as well as follow-up data, if applicable.

☒ The study involves data that exist at the time of IRB submission **and** data that will be generated after IRB submission. Include this activity in the Methods section.
Examples

- The study plans to conduct a retrospective chart review and ask subjects to complete a questionnaire.
- The study plans to include subjects previously diagnosed with a specific disease and add newly diagnosed subjects in the future.

☐ The study will use data that have been collected under another IRB protocol. Include in the Methods section and enter the IRB number from which the research material will be obtained. *When appropriate, note when subjects have provided consent for future use of their data and/or specimens as described in this protocol*.

Enter one IRB number per line, add more lines as needed

☐ Data ☐ Specimens ☐ Data & Specimens _____

☐ Data ☐ Specimens ☐ Data & Specimens _____

☐ Data ☐ Specimens ☐ Data & Specimens _____

Data Analysis

Power analyses may not be appropriate if this is a feasibility or pilot study, but end-point analysis plans are always appropriate even if only exploratory. Provide all information requested below, or provide justification if not including all of the information.

Power Statement: We anticipate 25 participants to be adequate to identify the heterogeneity of outcomes that matter to people living with T2D. Sampling will be guided by the principle of response saturation. We will use *stratified purposeful sampling* to intentionally select participants representing predefined demographic (age, sex, race/ethnicity, insurance type) and clinical (HbA_{1c} level, T2D drugs used) characteristics.[11,12] Analysis (i.e. generation of the individually weighted composite outcome) will occur concurrent with data collection, and we anticipate reaching saturation in the variability of patient preference rankings of the various outcomes with no more than 25 participants.[13-15]

Data Analysis Plan:

Quantitative analysis of the participatory ranking exercise: Each participant's responses will be used to generate an individualized weighted composite outcome. We will consider a utility function that is a weighted combination of the individual endpoints specified by each participant, with weights determined by the patient's preferences as proposed in Butler et al.[16] The patient's preference weights will be scaled to sum to 1, such that the patient-specific outcome will be a convex combination of the eligible outcomes. With the weighted utility function, we then define potentially counterfactual treatment effects, defined as the expected utility function under the four different treatment exposures. The optimal individual treatment rule is defined as in Qian and Murphy[17] as the treatment that maximizes the expected utility function.

We will then use the composite outcomes generated using the participator ranking exercise (see above) in a study that uses de-identified data within OLDW. Specifically, we will emulate a target trial comparing the effectiveness of GLP1-RA, SGLT2i, DPP-4i, and SU classes with respect to each of these patient-specified composite outcomes. The target parameter will be the six pairwise differences in the personalized utility function between the drug classes. We provide an overview of that analysis below, though all analyses using

OLDW data have been previously deemed exempt from IRB review as all data is statistically de-identified by Optumlabs prior to being made available to researchers.

Primary analyses will be conducted under an intention-to-treat approach, modified to define "intention" as a patient having two consecutive fills of the assigned drug. All patients will continue follow-up and complete all outcome assessments until the planned conclusion of the study or end of observation in OLDW.

We will estimate multinomial propensity score models using data from 12 months prior to the index date to account for the multiple treatments.[18] Predictions from the estimated propensity score model can be used for inverse probability weighting (IPW) to emulate baseline randomization. We will allow for flexible modeling strategies with controls to assess for overfitting. Our proposal is to use the super learner framework[19] with a diverse set of individual prediction algorithms included in the ensemble. The super learner framework allows for flexible estimation of an ensemble predictor with the theoretical properties of cross-validation for model selection to control for overfitting. The framework has been demonstrated to outperform individual algorithms for the estimation of propensity scores.[20,21] The distribution of the predicted propensity scores will be evaluated for evidence of violations of the positivity assumption. For moderate violations of positivity, the propensity scores will be truncated at 0.01 and 0.99 to control inflation of the standard errors. Our intention is to use the entire distribution of OLDW as our population of interest, but if more extreme positivity violations are observed, the reference population will be narrowed to a realistic treatment intervention population.

The proposed IPW weight with multinomial treatment groups is unstablized, as our primary estimand is the population average treatment effect.[22] If weights have extreme values when evaluating the distributions, then the stabilized version using the denominator proposed in Yoshida[23] for the optimal matching stabilized weights with multiple treatments will be used as a sensitivity analysis to avoid large confidence intervals for the effect estimates. The final propensity scores will then be incorporated as IPWs into a time-to-event model. If the proportional hazards assumptions hold, this will be a Cox model; otherwise, we will use Akaike Information Criteria to identify the best parametric time-to-event specification for the data. For the final model, we will report model diagnostic and performance statistics, and either hazard or time ratios with 95% confidence intervals. A single model with all treatment groups will be estimated and the omnibus test for at least one treatment group having a different cumulative incidence rate will be evaluated. If that test is statistically significant, then each pairwise comparison will be tested and pairwise effects estimated.

The targeted learning algorithm for the causal hazard rate as defined in Moore and van der Laan[24] will be implemented to estimate the four treatment hazard rates and the corresponding hazard ratios or time ratios, according to the model specification, will be estimated for each of the 6 pairwise comparisons between treatment groups. A single overall omnibus test for the four hazard rates being equal to each other versus at least one different will be conducted at the 0.05 significance level, with pairwise comparison p-values adjusted using the Holm closed sequential multiple testing procedure.[25,26]

Predefined subgroup analyses for heterogeneity of treatment effects (HTE) will be evaluated for the following characteristics: sex (female vs. male), race/ethnicity (Black, Hispanic, Asian, other vs. White), and age (45-64, 65-74, ≥75 vs. <45 years). While there is no definitive evidence of HTE by these characteristics, the incidence of CVD and kidney disease outcomes are known to vary[27-30] and may be mediated by different responses to therapeutic interventions. We will also assess for HTE as a function of different levels of underlying CVD risk, as the examined drugs may be more effective at reducing the risks of primary and secondary endpoints among patients at higher risk for these events or with certain CVD risk factors already present.[31-35 36-39] The primary model will be used to estimate treatment effects across subgroups, with the targeted parameter for the subgroup analysis being the difference in the predicted treatment effect for each pairwise comparison between the

subgroups.[40] We will estimate multivariable effect prediction models[41,42] to assess for HTE between the predefined subgroups by incorporating specific parameters for the statistical interaction effect between the treatment and subgroup variables.[36,37] Propensity score weights will be used to emulate randomization according to the target trial framework. Internal cross-validation will be used to control for overfitting of the prediction model.[43] Adjustment for multiple testing will be done accounting for the total number of predefined HTE analyses.

A secondary analysis will be conducted under an as-treated approach, whereby patients will be censored upon discontinuation of the assigned treatment (no fills after 30 days beyond the dispensed pill count) or addition of another drug (i.e. another 2nd-line drug or insulin). To adjust for the likely informative censoring post-baseline, time-dependent propensity score models will be estimated and applied to dynamically reweight the population and account for time-varying confounding and informative loss to follow-up. The time-dependent propensity score models will be based on all prior covariate information up to that time point. This will address two concerns with use of observational methods for comparative effectiveness, immortal time bias[44] and time-lag bias,[45] as we will allow for time-dependent exposures and reweighting on time-conditional propensity scores.[46]

Qualitative analysis of the accompanying interview: Interviews will be transcribed and analyzed according to a general thematic analysis approach (Braun & Clark, 2012) to identify patterns of meaning across patient interviews.

MAYO CLINIC

REFERENCES

1. CDC. National Diabetes Statistics Report, 2020. In. Atlanta, Georgia, U.S.A.: U.S. Department of Health and Human Services, Centers for Disease Control and Prevention; 2020.
2. American Diabetes Association Standards of Medical Care in Diabetes—2020. Section 9. Pharmacologic Approaches to Glycemic Treatment. *Diabetes Care.* 2020;43(Supplement 1):S98-S110.
3. NICE. National Institute for Health and Care Excellence Pathways: Managing Blood Glucose in Aults with Type 2 Diabetes. National Institute for Health and Care Excellence. https://pathways.nice.org.uk/pathways/type-2-diabetes-in-adults. Published 2019. Updated March 26, 2019. Accessed April 23, 2019.
4. Garber AJ, Abrahamson MJ, Barzilay JI, et al. Consensus Statement By The American Association Of Clinical Endocrinologists And American College Of Endocrinology On The Comprehensive Type 2 Diabetes Management Algorithm – 2019 Executive Summary. *Endocrine Practice.* 2019;25(1):69-100.
5. Conlin PR, Colburn J, Aron D, Pries R, Tschanz MP, Pogach L. Synopsis of the 2017 U.S. Department of Veterans Affairs/U.S. Department of Defense Clinical Practice Guideline: Management of type 2 diabetes mellitus. *Annals of Internal Medicine.* 2017;167(9):655-663.
6. Baena-Díez JM, Peñafiel J, Subirana I, et al. Risk of Cause-Specific Death in Individuals With Diabetes: A Competing Risks Analysis. *Diabetes Care.* 2016;39(11):1987-1995.
7. Diabetes Mellitus, Fasting Glucose, and Risk of Cause-Specific Death. *New England Journal of Medicine.* 2011;364(9):829-841.
8. Tancredi M, Rosengren A, Svensson A-M, et al. Excess Mortality among Persons with Type 2 Diabetes. *New England Journal of Medicine.* 2015;373(18):1720-1732.
9. Afkarian M, Sachs MC, Kestenbaum B, et al. Kidney disease and increased mortality risk in type 2 diabetes. *J Am Soc Nephrol.* 2013;24(2):302-308.
10. Janssen IM, Gerhardus A, Schroer-Gunther MA, Scheibler F. A descriptive review on methods to prioritize outcomes in a health care context. *Health Expect.* 2015;18(6):1873-1893.
11. Luborsky MR, Rubinstein RL. Sampling in Qualitative Research: Rationale, Issues, and Methods. *Research on aging.* 1995;17(1):89-113.
12. Palinkas LA, Horwitz SM, Green CA, Wisdom JP, Duan N, Hoagwood K. Purposeful Sampling for Qualitative Data Collection and Analysis in Mixed Method Implementation Research. *Adm Policy Ment Health.* 2015;42(5):533-544.
13. Sargeant J. Qualitative Research Part II: Participants, Analysis, and Quality Assurance. *Journal of Graduate Medical Education.* 2012;4(1):1-3.
14. Guest G, Bunce A, Johnson L. How Many Interviews Are Enough? An Experiment with Data Saturation and Variability. *Field Methods.* 2006;18(1):59-82.
15. Saunders B, Sim J, Kingstone T, et al. Saturation in qualitative research: exploring its conceptualization and operationalization. *Qual Quant.* 2018;52(4):1893-1907.
16. Butler EL, Laber EB, Davis SM, Kosorok MR. Incorporating Patient Preferences into Estimation of Optimal Individualized Treatment Rules. *Biometrics.* 2018;74(1):18-26.
17. Qian M, Murphy SA. Performance Guarantees for Individualized Treatment Rules. *Ann Stat.* 2011;39(2):1180-1210.
18. McCaffrey DF, Griffin BA, Almirall D, Slaughter ME, Ramchand R, Burgette LF. A tutorial on propensity score estimation for multiple treatments using generalized boosted models. *Stat Med.* 2013;32(19):3388-3414.

19. van der Laan MJ, Polley EC, Hubbard AE. Super learner. *Stat Appl Genet Mol Biol.* 2007;6:Article 25.
20. Sofrygin O, Zhu Z, Schmittdiel JA, et al. Targeted learning with daily EHR data. *Stat Med.* 2019;38(16):3073-3090.
21. Kyle RP, Moodie EEM, Klein MB, Abrahamowicz M. Evaluating Flexible Modeling of Continuous Covariates in Inverse-Weighted Estimators. *Am J Epidemiol.* 2019;188(6):1181-1191.
22. Li F, Li F. Propensity score weighting for causal inference with multiple treatments. *Ann Appl Stat.* 2019;13(4):2389-2415.
23. Yoshida K, Hernández-Díaz S, Solomon DH, et al. Matching Weights to Simultaneously Compare Three Treatment Groups: Comparison to Three-way Matching. *Epidemiology (Cambridge, Mass).* 2017;28(3):387-395.
24. Moore KL, Van der Laan MJ. Application of time-to-event methods in the assessment of safety in clinical trials. In: Peace KE, ed. *Design and Analysis of Clinical Trials with Time-to-Event Endpoints.* Chapman and Hall/CRC; Biostatistics Series; 2009:455-482.
25. Takahashi K, Serruys PW, Fuster V, et al. Redevelopment and validation of the SYNTAX score II to individualise decision making between percutaneous and surgical revascularisation in patients with complex coronary artery disease: secondary analysis of the multicentre randomised controlled SYNTAXES trial with external cohort validation. *Lancet.* 2020;396(10260):1399-1412.
26. Bress AP, Greene T, Derington CG, et al. Patient Selection for Intensive Blood Pressure Management Based on Benefit and Adverse Events. *J Am Coll Cardiol.* 2021;77(16):1977-1990.
27. Karter AJ, Lipska KJ, O'Connor PJ, et al. High rates of severe hypoglycemia among African American patients with diabetes: the surveillance, prevention, and Management of Diabetes Mellitus (SUPREME-DM) network. *Journal of diabetes and its complications.* 2017;31(5):869-873.
28. Chang MH, Moonesinghe R, Athar HM, Truman BI. Trends in Disparity by Sex and Race/Ethnicity for the Leading Causes of Death in the United States-1999-2010. *J Public Health Manag Pract.* 2016;22 Suppl 1:S13-24.
29. Kautzky-Willer A, Harreiter J, Pacini G. Sex and Gender Differences in Risk, Pathophysiology and Complications of Type 2 Diabetes Mellitus. *Endocrine Reviews.* 2016;37(3):278-316.
30. Prospective Studies C, Asia Pacific Cohort Studies C. Sex-specific relevance of diabetes to occlusive vascular and other mortality: a collaborative meta-analysis of individual data from 980 793 adults from 68 prospective studies. *Lancet Diabetes Endocrinol.* 2018;6(7):538-546.
31. Tsapas A, Avgerinos I, Karagiannis T, et al. Comparative Effectiveness of Glucose-Lowering Drugs for Type 2 Diabetes. *Annals of Internal Medicine.* 2020;173:278-286.
32. Zheng SL, Roddick AJ, Aghar-Jaffar R, et al. Association Between Use of Sodium-Glucose Cotransporter 2 Inhibitors, Glucagon-like Peptide 1 Agonists, and Dipeptidyl Peptidase 4 Inhibitors With All-Cause Mortality in Patients With Type 2 Diabetes: A Systematic Review and Meta-analysis. *JAMA.* 2018;319(15):1580-1591.
33. Zelniker TA, Wiviott SD, Raz I, et al. Comparison of the Effects of Glucagon-Like Peptide Receptor Agonists and Sodium-Glucose Cotransporter 2 Inhibitors for Prevention of Major Adverse Cardiovascular and Renal Outcomes in Type 2 Diabetes Mellitus. *Circulation.* 2019;139(17):2022-2031.
34. Zelniker TA, Wiviott SD, Raz I, et al. SGLT2 inhibitors for primary and secondary prevention of cardiovascular and renal outcomes in type 2 diabetes: a systematic review and meta-analysis of cardiovascular outcome trials. *Lancet.* 2019;393(10166):31-39.
35. Selvin E, Bolen S, Yeh HC, et al. Cardiovascular outcomes in trials of oral diabetes medications: a systematic review. *Arch Intern Med.* 2008;168(19):2070-2080.

36. Kent DM, van Klaveren D, Paulus JK, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement: Explanation and Elaboration. *Ann Intern Med.* 2020;172(1):W1-W25.
37. Kent DM, Paulus JK, van Klaveren D, et al. The Predictive Approaches to Treatment effect Heterogeneity (PATH) Statement. *Ann Intern Med.* 2020;172(1):35-45.
38. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ.* 2018;363:k4245.
39. Heerspink HJL, Stefansson BV, Correa-Rotter R, et al. Dapagliflozin in Patients with Chronic Kidney Disease. *N Engl J Med.* 2020;383(15):1436-1446.
40. VanderWeele TJ, Luedtke AR, van der Laan MJ, Kessler RC. Selecting Optimal Subgroups for Treatment Using Many Covariates. *Epidemiology (Cambridge, Mass).* 2019;30(3):334-341.
41. Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *Bmj.* 2018;363.
42. Varadhan R, Segal JB, Boyd CM, Wu AW, Weiss CO. A framework for the analysis of heterogeneity of treatment effect in patient-centered outcomes research. *Journal of clinical epidemiology.* 2013;66(8):818-825.
43. Arlot S, Celisse A. A survey of cross-validation procedures for model selection. *Statist Surv.* 2010;4:40-79.
44. Suissa S. Immortal time bias in observational studies of drug effects. *Pharmacoepidemiol Drug Saf.* 2007;16(3):241-249.
45. Suissa S. Lower Risk of Death With SGLT2 Inhibitors in Observational Studies: Real or Bias? *Diabetes Care.* 2018;41(1):6-10.
46. Suissa S, Moodie EE, Dell'Aniello S. Prevalent new-user cohort designs for comparative drug effect studies by time-conditional propensity scores. *Pharmacoepidemiol Drug Saf.* 2017;26(4):459-468.