**Appendix 1: List of institutions and affiliated dialysis organizations in the TwoPlus Trial**

| Institution | Affiliated Dialysis Organization |
|---|---|
| Emory University | University-owned |
| Johns Hopkins University School of Medicine | Fresenius Medical Care |
| Massachusetts General Hospital | Fresenius Medical Care |
| Northwell | DaVita Kidney Care |
| Renal Research Institute | Fresenius Medical Care |
| University of Florida Jacksonville | Fresenius Medical Care; Dialysis Clinic Incorporated |
| University of Mississippi Medical Center | University-owned |
| University of North Carolina | Fresenius Medical Care |
| University of Virginia | University-owned |
| Wake Forest University School of Medicine | University-owned |

**Appendix 2: Footnotes for the eligibility criteria described in Table 1**

Notes Regarding Eligibility Criteria

[1]: The decision for study eligibility (i.e., eligibility for twice-weekly HD) should be made by the treating nephrologist in conjunction with the Site Investigators.

[2]: The time when the patient is approached for potential study participation refers to the date when the patient is approached for getting an informed consent for study participation (i.e., the date when a member of the study team discusses with the patient about the study and offers study participation). All efforts must be made to assess eligibility and recruit as early as possible after HD initiation, ideally within 1-2 weeks of HD commencement. For patients who are anticipated to be started on maintenance, in-center HD, decision of HD initiation is made by the treating providers.

[3]: Both parameters of residual kidney function (kidney urea clearance and urine volume) must be met to be eligible for study participation.

[4]: Examples of high-volume ultrafiltration requirements include interdialytic weight gain ≥ 4% of target weight; post-HD residual weight ≥5% of target weight. Note that these are not absolute requirements.

[5]: If abnormal values, re-assess at a later date (e.g., after 12 HD sessions, sooner, or later) provided patient would not have been on HD for >12 weeks (i.e., first date of HD was not >84 days prior) at the time patient will be approached for potential study participation; and target weight was dynamically re-assessed and gauged by the Site Investigators and treating providers.

[6]: Examples given: urinary incontinence; severe dementia; history of medical non-adherence that, in the opinion of the Site Investigators and/or treating providers, precludes safe study participation; a medical condition that, in the opinion of the Site Investigators and/or treating providers, would jeopardize the safety of the participant; or unable or unwilling to give informed consent for study participation.

[7]: If results of baseline residual kidney function levels (calculated on baseline timed urine collection) are available through chart review, they are taken into consideration for eligibility criteria if the urine collection occurred within 2 weeks before the prescreening date. If more than 2 weeks have elapsed, then recruitment should follow scenario a), i.e., a new baseline timed urine collection will be obtained at the time of screening, after informed consent for study participation is obtained.

[8]: In situations of inconclusive results of screening RKF parameters following the first RKF testing, the investigators can repeat RKF testing, especially if there are reasons to suspect the first timed urine collection was inaccurate.

[9]: We define women of childbearing capacity as women ≤55 years old who do not have a history of hysterectomy.

Kidney dysfunction requiring dialysis (KDRD) connotes the same concept as end-stage kidney disease (ESKD) or end-stage renal disease (ESRD) or chronic kidney disease stage 5 on dialysis (CKD5D).

The decision to initiate chronic, in-center HD is made by the treating providers, independent of this study.

Target weight connotes the same concept as Dry Weight; the terminology is used interchangeably.

Calculation of Inter-dialytic Weight Gain (IDWG) as % of Target Weight:

$$IDWG = [(\text{pre-HD weight} - \text{target weight})/\text{target weight}] \times 100$$

Calculation of Residual Weight as % of Target Weight:

Residual weight = [(post-HD weight – target weight)/target weight] x 100

Notes regarding the number of dialysis sessions: "Has received ≤36 sessions of intermittent HD or has been on dialysis for ≤12 weeks at the time patient is approached for potential study participation."

- We recognize patients may undergo different types of dialysis that are not interchangeable in terms of the number of dialysis treatments. In these types of situations, for example, when patients underwent continuous renal replacement therapy (CRRT) and/or NxStage HD, besides conventional intermittent HD, the dialysis duration of '≤12 weeks', rather than the dialysis duration of '≤36 sessions of intermittent HD', will be used to gauge eligibility.
- We recognize patients may have attempts of intermittent HD that are incomplete due to vascular access infiltration or catheter malfunction. Incomplete HD sessions of this nature may not need to be included in the tally of '36 HD sessions'.

Notes regarding the requirement or anticipated requirement of high-volume ultrafiltration[4,5]

It is acknowledged that the target weight for patients undergoing HD is not a fixed value and can vary depending on individual circumstances. This variability occurs because the determination of target weight is not based on precise scientific measurements. In some cases, the target weight may be over- or under-estimated, leading to inaccuracies in calculations such as interdialytic weight gain and residual weight. Therefore, meeting the criteria of interdialytic weight gain ≥ 4% of target weight and/or post-HD residual weight ≥ 5% of target weight is not a strict requirement. The target weight should continually be re-assessed by the treating team according to usual care. The decision for study eligibility should be made based on a clinical assessment conducted by either the study investigator or a member of the treating team, such as a nephrologist or advanced practice provider. This assessment should take into consideration the patient's stability in terms of ultrafiltration needs, as well as the potential for adjusting the frequency of hemodialysis sessions.

Notes regarding Acute Kidney Injury (AKI) pending ESRD transition:

Patients started on intermittent HD with a diagnosis of Acute Kidney Injury (AKI) or Acute Renal Failure (ARF) may be approached for study participation only under the following conditions: i) if the Site PI determines, based on clinical evidence, that the patient will not have recovery of kidney function, will require chronic HD, and will be diagnosed with End-Stage Renal Disease (ESRD) or End-Stage Kidney Disease (ESKD); or ii) after the patient is diagnosed with ESRD or ESKD. The medical decision for a diagnosis of ESRD or ESKD must be made independent of the TwoPlus study. In these cases, the total number of HD treatments or the total duration of dialysis will be calculated from the first date of dialysis administration with the AKI diagnosis. The patient can be approached for study participation only if the eligibility criteria of dialysis vintage are met—i.e., patient has received ≤36 sessions of intermittent HD or has been on dialysis for ≤12 weeks.

Notes regarding pregnancy-related eligibility assessment:

Definition: woman of childbearing capacity

A woman of childbearing capacity is a person of female sex, age ≤55 years, and with presence of an anatomical uterus.

For women aged 55 or under, the sole circumstance in which they are considered to lack childbearing capacity is if the Electronic Medical Records (EMR) within the healthcare system document a history of undergoing a total hysterectomy.

A history of total hysterectomy must be evidenced in the healthcare system EMR. Scenario when the history of hysterectomy is reported by the patient but not identified in the healthcare system, EMR will be considered childbearing capacity.

Women aged ≤55 years with either i) reported amenorrhea, or ii) tubal ligation, or iii) intrauterine device, or iv) on contraceptive methods are considered of childbearing capacity.

## Appendix 3. Criteria for Progression from Twice-weekly to Thrice-weekly HD[1]

| |
|---|
| **Residual kidney function criteria[2]** |
| 1. Kidney urea clearance <2.0 mL/min |
| 2. Urine volume <500 mL/24 h, despite optimized diuretic regimen |
| **Clearance criteria by blood tests[3]** |
| 3. Pre-HD serum K ≥5.8 mEq/L, with or without EKG changes of hyperkalemia, despite treatment with oral K binder |
| 4. Pre-HD serum Na ≤125 mEq/L |
| 5. Pre-HD serum bicarbonate level ≤17 mEq/L, despite optimized treatment with oral sodium bicarbonate |
| 6. Unable to attain spKt/V ≥1.20 |
| 7. Unable to attain total stdKt/V (kidney stdKt/V + dialysis stdKt/V) ≥2.10 |
| **Volume management criteria** |
| 8. Inter-dialytic weight gain ≥ 4% of target weight, despite optimized diuretic regimen |
| 9. Ultrafiltration rate ≥13 mL/kg/h, despite optimized diuretic regimen |
| 10. Post-HD residual weight ≥5% of target weight, despite optimized diuretic regimen |
| 11. Uncontrolled heart failure, which, in the opinion of the treating nephrologist, warrants more frequent HD |
| 12. Uncontrolled hypertension, which, in the opinion of the treating nephrologist, warrants more frequent HD |
| **Other clinical criteria** |
| 13. EKG abnormalities which, in the opinion of the treating nephrologist, are due to electrolyte and/or acid-base disequilibrium and require more frequent HD |
| 14. Clinical event that requires ≥1 unplanned HD treatment for its resolution, believed to be related to twice-weekly HD |
| 15. Clinical condition, which, in the opinion of the treating nephrologist, warrants more frequent HD |
| 16. Non-adherence to timed urine collection for ≥2 consecutive occasions when timed urine collection was recommended by the treating providers and/or Site Investigators |
| 17. Intravenous antibiotic administration required thrice-weekly with HD[4] |
| 18. New condition rendering the patient with inability to perform timed urine collection (e.g., stroke, urinary incontinence) |

**Abbreviations:** EKG, electrocardiogram; HD, hemodialysis; spKt/V, single pool Kt/V; stdKt/V, standard Kt/V.

**Criterial for progression footnotes**

[1]: The decision to transition from twice- to thrice-weekly HD is made by the treating nephrologist in conjunction with the Site Investigator, not by a study coordinator. No criterion, taken in isolation, is an absolute indication for transitioning from twice- to thrice-weekly HD. Each criterion must be judged in the overall clinical context for each individual to decide on medical necessity and timing of transition from twice- to thrice-weekly HD.

[2]: Either one parameter of residual kidney function (i.e., either kidney urea clearance or urine volume) or both parameters can qualify the patient to transition from twice- to thrice-weekly HD.

[3:] On one or two determinations among 3 to 18.

[4:] Residual kidney function may be re-assessed within 2 weeks of antibiotic treatment completion to evaluate whether conversion back to twice-weekly HD is acceptable, provided no other indication exists for thrice-weekly HD.

**Appendix 4: Detailed description of the primary, secondary, and exploratory outcomes**

## Primary outcome

**Incidence rate ratio of all-cause emergency department visits, hospitalizations, or death**
*Description*: The total safety event incidence rate, in each treatment group, will comprise all-cause emergency department (ED) visits not leading to a hospitalization, all-cause hospitalizations, and all-cause mortality. This rate will be calculated as the total number of safety events divided by the total number of person-years observed in each treatment group. A participant could have recurring ED visits, recurring hospitalizations, and/or die during the study, and all events will be included in this rate.
*Time Frame*: From date of randomization to an end-of-study event.

## Secondary outcomes

### *Main Secondary Outcomes*

**Illness Intrusiveness Rating Scale (IIRS) score**
*Description*: Difference in the change from baseline observed between patients randomized to CHD, compared to those randomized to CMIHD. Comparisons will be made a) between the two treatment groups for the duration of the study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within the conventional HD group before and after transition observed in clinically-matched incremental HD group. The time point of transition represents the time point at which patient participants in the clinically-matched incremental HD group are switched from twice-weekly to thrice-weekly HD according to the Criteria for Progression.
*Time Frame*: The IIRS questionnaire will be administered monthly, starting at baseline pre-randomization to the end-of-study event.

**Change in urine volume**
*Description*: Timed urine collection will be obtained monthly or no less often than every 3 months. Difference in the change from baseline observed between patients randomized to CHD, and trajectory of urine volume, in mL/24 hour, and the average urine volume at baseline and at the end-of-study event will be summarized to help with interpretation of results. Change from baseline in urine volume will be compared a) between treatment groups, from baseline to end-of-study; b) within clinically-matched incremental HD group, before and after transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after transition observed in clinically-matched incremental HD group. Time from baseline to urine volume is <100 mL/24 hour will be compared between treatment groups. The time point of transition represents the time point at which patient participants in the clinically-matched incremental HD group are switched from twice-weekly to thrice-weekly HD according to the Criteria for Progression.
*Time Frame*: For continuous outcome analysis: from date of randomization to end-of-study event. For binary outcome analysis: from date of randomization to time point of documented urine output <100 mL/24 hour based on timed urine collection.

**Change in renal urea clearance**
*Description*: Timed urine collection will be obtained monthly or no less often than every 3 months. Difference in the change from baseline observed between patients randomized to CHD, and trajectory of renal urea clearance, in mL/min and mL/min/1.73m$^2$, and the average urea clearance at baseline and at the end-of-study event will be summarized to help with interpretation of results.
Change from baselinerenal urea clearance will be compared a) between treatment groups, from baseline to end-of-study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after transition in clinically-matched incremental HD group. The time point of transition represents the time point at which patient participants in the clinically-matched incremental HD group are switched from twice-weekly to thrice-weekly HD according to the established Criteria for Progression. Rate of change in renal urea clearance (to be expressed in rate per month) will also be compared between treatment groups.
*Time Frame*: From the date of randomization to end-of-study event.

**Change in kidney creatinine clearance**

*Description*: Timed urine collection will be obtained monthly or no less often than every 3 months. Difference in the change from baseline observed between patients randomized to CHD, and trajectory of renal creatinine clearance, in mL/min and mL/min/1.73m$^2$, and the average renal creatinine clearance at baseline and at the end-of-study event will be summarized to help with interpretation of results.
Change from baseline renal creatinine clearance, in mL/min and in mL/min/1.73m$^2$, will be calculated at the end of study; Change from baseline in renal creatinine clearance will be compared a) between treatment groups, from baseline to end-of-study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after transition in clinically-matched incremental HD group. The time point of transition represents the time point at which patient participants in the clinically-matched incremental HD group are switched from twice-weekly to thrice-weekly HD according to *Criteria for Progression*. Rate of change in renal creatinine clearance, in rate per month, will also be compared between treatment groups.
*Time Frame*: From the date of randomization to end-of-study event.

**EuroQOL-5 Dimensions-5 Level (EQ-5D-5L) score**
*Description*: Baseline-adjusted EQ-5D-5L score will be compared a) between treatment groups at the time points of questionnaire administration for the duration of the study; b) within clinically-matched incremental HD group, before and after transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after transition in clinically-matched incremental HD group. The time of transition represents the time point at which patient participants in the clinically-matched incremental HD group are switched from twice-weekly to thrice-weekly HD according to Criteria for Progression.
*Time Frame*: The EQ-5D-5L questionnaire will be administered at baseline pre-randomization, and at months 6, 12, 18, and 24.

***Additional Secondary Outcomes***

**Hospital-free days per 100 patient-days**
*Description*: The duration of hospital-free days will be determined for each patient and calculated in the whole cohort per 100 patient-days. It is anticipated that we will observe an array of hospitalization events for each participant, ranging from no hospitalization to frequent hospitalizations per study period. All days of ED visits, hospitalizations (inpatient and observation status), and long-term acute-care hospital (LTACH), per participant, will be analyzed. A period of hospitalization will be computed from the date of hospital admission to the date of discharge. Hospital-free days may include one or more discrete time segments of non-hospitalization between hospitalization periods. The total hospital-free days per patient participant will be calculated as total study days – hospitalization days. Hospital-free days will be normalized per 100 patient-days.
*Time Frame*: From date of randomization to end-of-study event.

**Urine screen**
*Description*: Baseline-adjusted urine screen score will be compared between treatment groups at 6, 12, 18, and 24, using an average of answers obtained during monthly questionnaire administration.
*Time Frame*: From the date of randomization to an end-of-study event.

**Time to recover from hemodialysis (HD)**
*Description*: The question 'How long does it take to recover from a dialysis session?' will be administered to all patient participants monthly. Four answer categories will be offered: <2, 2–6, 7–12, and >12 hours. Comparisons will be made a) between treatment groups at the time points of questionnaire administration for the duration of the study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after the average time point of transition observed in clinically-matched incremental HD group.
*Time Frame*: From the date of randomization to an end-of-study event.

**SONG-HD Fatigue score**
*Description*: The SONG-HD Fatigue scale consists of three items that assess the effect of fatigue on life participation, tiredness, and energy level. These dimensions are assessed on a four-point Likert scale indicating increasing severity, ranging from zero (not at all) to three (severely). An overall score for fatigue is obtained by summing the responses across the three questions, resulting in a scale ranging from zero (no fatigue) to nine (maximum fatigue). Patient participants will be asked to respond on the basis of their

experience of fatigue in the past week. Comparisons will be made between treatment groups at the time points of questionnaire administration for the duration of the study.
*Time Frame*: Baseline pre-randomization; and months 6, 12, 18, and 24

**Trail Making Test B score**
*Description*: Trail Making Test B evaluates the ability to visually search, sustain attention, and perform cognitive shifting as the activity is completed. The subject draws lines, connecting 25 circles that contain numbers from 1 to 13 and letters from A to L. The subjects must draw lines alternating from number to letter. The total time to complete the task is logged. Errors are not counted, but the subject is alerted to mistakes made. They are instructed to correct them, which increases the amount of time needed to complete the task.
*Time Frame*: Baseline pre-randomization; and months 6, 12, 18, and 24.

**Employment status**
*Description*: The Employment Survey will ask the patient participant if he/she is employed for pay at a job or business at the time of survey administration, and in the past, with the last date of employment. We will calculate adjusted odds ratios (ORs) and 95% confidence intervals for 'lost employment', 'return to employment', and 'new employment', including the following mutually adjusted variables: age at enrollment; health insurance type; education completed; and baseline comorbidity index. Comparisons will be made between treatment groups at the time points of assessment for the duration of the study.
*Time Frame*: Baseline pre-randomization; and months 6, 12, 18, and 24.

**Volume management: interdialytic weight gain, ultrafiltration rate, and residual weight**
*Description*: Interdialytic weight gain will be calculated as [(pre-HD weight – post-HD weight; weight removed with each HD will be calculated as pre-HD weight – post-HD weight; ultrafiltration rate will be calculated as weight removed/target weight/treatment time (in mL/kg/h); residual weight will be calculated as [(post-HD weight-target weight)/target weight]x100 (in kg and as % of target weight). A monthly average for these variables will be calculated based on data observed with each HD treatment during each month. These outcomes will be analyzed as continuous outcomes or binary outcomes (proportion of patients with ≥6% of target weight for interdialytic weight; ≥13 mL/kg/h for ultrafiltration rate, and ≥2 kg above target weight for residual weight). Comparisons will be made a) between treatment groups for the duration of the study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after the average time point of transition observed in clinically-matched incremental HD group. The time point of transition represents the time point at which patient participants in the clinically-matched incremental HD group are switched from twice-weekly to thrice-weekly HD according to *Criteria for Progression*.
*Time Frame*: From the date of randomization to end-of-study event.

**Cardiovascular parameters: pre- and post-HD blood pressure and heart rate; intradialytic peak and nadir blood pressure and heart rate; and number of antihypertensive medications**
*Description*: Average pre-HD blood pressure and heart rate, post-HD blood pressure and heart rate, intradialytic peak blood pressure and heart rate, intradialytic nadir blood pressure and heart rate, and number of prescribed antihypertensive medications will be calculated monthly based on data observed with each HD treatment during the month and medication list review. Each outcome will be adjusted for baseline data. Comparisons will be made a) between treatment groups for the duration of the study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after the average time point of transition observed in clinically-matched incremental HD group.
*Time Frame*: From the date of randomization to end-of-study event.

**Urea clearance parameters: spKt/V, renal stdKt/V, dialysis stdKt/V and total stdKt/V**
*Description*: Monthly spKt/V obtained at outpatient dialysis units will be collected. Renal stdKt/V and dialysis stdKt/V will be calculated monthly with the use of Solute Solver equation. A centralized database system will provide these calculations. Parameters required to make these calculations include spKt/V, urine volume, renal urea clearance, pre-HD and post-HD blood urea nitrogen, and ultrafiltration on the day timed urine collection is submitted, duration of timed urine collection, HD frequency, and HD duration for each session. Total stdKt/V denotes the sum of renal stdKt/V and dialysis stdKt/V.
*Time Frame*: From the date of randomization to an end-of-study event.

**Biochemical parameters: pre-HD serum potassium (K), serum sodium (Na) and serum bicarbonate (TCO2)**
*Description*: Monthly pre-HD serum K, serum Na, and serum TCO2 obtained at outpatient dialysis units will be collected. When more than one measurement is obtained during a month, an average of all measurements will be collected. These parameters will be analyzed as continuous and binary variables. Comparisons will be made a) between treatment groups at the time points of questionnaire administration for the duration of the study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after the average time point of transition observed in clinically-matched incremental HD group.
*Time Frame*: From the date of randomization to an end-of-study event.

**Parameters of anemia management: hemoglobin, ferritin, transferrin saturation, epogen dose, and intravenous iron dose**
*Description*: Monthly pre-HD hemoglobin, ferritin, transferrin saturation obtained at outpatient dialysis units will be collected. Medications administered at dialysis will be collected monthly as the average weekly dose of active vitamin D analogue (in mcq/wk hectorol dose equivalent), erythropoietin stimulating agent (in units/wk epogen dose equivalent) and intravenous iron (in mg/wk iron sucrose dose equivalent), calculated based on the dose/week administered over the 4 weeks of monthly review. Comparisons will be made a) between treatment groups at the time points of questionnaire administration for the duration of the study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after the average time point of transition observed in clinically-matched incremental HD group.
*Time Frame*: From date of randomization to end-of-study event.

**Parameters of mineral metabolism management: phosphorus, calcium, intact parathyroid hormone (iPTH), active vitamin D analogue dose, and calcimimetic dose**
*Description*: Monthly pre-HD serum phosphorus, calcium, an iPTH levels obtained at outpatient dialysis units will be collected. Medications administered at dialysis will be collected monthly as the average weekly dose of active vitamin D analogue (in mcq/wk hectorol dose equivalent) and calcimimetic dose (in mg/wk cinacalcet dose equivalent), calculated based on the dose/week administered over the 4 weeks of monthly review. Comparisons will be made a) between treatment groups at the time points of questionnaire administration for the duration of the study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after the average time point of transition observed in clinically-matched incremental HD group.
*Time Frame*: From date of randomization to end-of-study event.

**Parameters of nutrition: serum albumin and normalized protein catabolic rate (nPCR)**
*Description*: Monthly pre-HD serum albumin and nPCR obtained at outpatient dialysis units will be collected. Comparisons will be made a) between treatment groups at the time points of questionnaire administration for the duration of the study; b) within clinically-matched incremental HD group, before and after the time point of transition from twice-weekly to thrice-weekly HD; and c) within conventional HD group before and after the average time point of transition observed in clinically-matched incremental HD group.
*Time Frame*: From date of randomization to end-of-study event.

**Incidence rate of all-cause and cause-specific emergency department (ED) visits**
*Description*: As a component of the primary composite outcome, we will compare the incidence rate of all-cause and cause-specific ED visits that did not lead to hospitalization. The cause of an ED visit will be categorized as non-vascular access-related or vascular access (central venous catheter or arteriovenous access)-related. Non-vascular access events will be subcategorized into cardiovascular, cerebrovascular, infectious, gastrointestinal, sudden death, malignancy-related, or other. Vascular access-related events will be subcategorized as due to infectious or non-infectious complications. The incidence rate of ED visits will be calculated as the total number of ED visits divided by the total number of person-years observed in each treatment group. A participant could have recurring ED visits, and all events will be included in this rate. Time to first ED visit will also be compared between the groups.
*Time Frame*: From date of randomization to end-of-study event.

**Incidence rate of all-cause and cause-specific hospitalizations**

*Description*: As a component of the primary composite outcome, we will compare the incidence rate of all-cause and cause-specific hospitalizations. The cause of a hospitalization will be categorized as non-vascular access related or vascular access (central venous catheter or arteriovenous access)-related. Non-vascular access events will be subcategorized into cardiovascular, cerebrovascular, infectious, gastrointestinal, sudden death, malignancy-related, or other. Vascular access-related events will be subcategorized as due to infectious or non-infectious complications. Hospitalization incidence rate will be calculated as the total number of hospitalizations divided by the total number of person-years observed in each treatment group. A participant could have hospitalizations, and all events will be included in this rate. Time to first hospitalization and total hospital days per patient-year will also be compared between the treatment groups.
*Time Frame*: From date of randomization to end-of-study event.

**All-cause and cause-specific death rate**
*Description*: As a component of primary composite outcome, we will compare all-cause and cause-specific death rates. The cause of death will be categorized as non-vascular access related or vascular access (central venous catheter or arteriovenous access)-related. Non-vascular access events will be subcategorized into cardiovascular, cerebrovascular, infectious, gastrointestinal, sudden death, malignancy-related, elective withdrawal of dialysis, sudden death outside the hospital, or other. Vascular access-related events will be subcategorized as due to infectious or non-infectious complications. Time to death will also be compared between the treatment groups.
*Time Frame*: From date of randomization to end-of-study event.

**Exploratory outcomes**

**Vascular access-related events: central venous catheter and/or arteriovenous access infectious and non-infectious complications**
*Description*: All vascular access procedures, both inpatient and outpatient procedures, will be tracked monthly. Infectious complications will include catheter-related bloodstream infection, sepsis or endocarditis, with or without other metastatic infections; arteriovenous access cellulitis or abscess or bacteremia, with or without other metastatic infections. Non-infectious complications will include arteriovenous access thrombosis. Vascular access procedures will include fistulogram, angioplasty, stenting, use of thrombolytics, removal of a permanent access and reason for removal, placement of a new permanent access; and catheter removal, exchange, or insertion.
*Time Frame*: From date of randomization to end-of-study event.

**Time to first major adverse cardiovascular or cerebrovascular event (MACCE)**
*Description: Time from randomization to first event of cardiac death, acute myocardial infarction (ST or non-ST elevation), coronary revascularization (angioplasty or coronary artery bypass surgery), hospitalization due to acute decompensated congestive heart failure, acute ischemic stroke (acute ischemic cerebrovascular accident [CVA]), or acute hemorrhagic stroke (acute hemorrhagic CVA).*
*Time Frame: From the date of randomization to the date of first MACCE.*

**Relationship between baseline covariates and the time point of transition from twice-weekly HD to thrice-weekly HD in the clinically-matched incremental HD group**
*Description*: The time point of transition from twice-weekly HD to thrice-weekly HD in the clinically-matched incremental HD group will be analyzed in a time-to-event analysis. Associations between baseline covariates of age, sex, race, ESKD etiology, residual kidney function level, and comorbidity index with interval time to transition from twice- to thrice-weekly HD will be explored among patient participants in the clinically-matched incremental HD group.
*Time Frame*: From the date of randomization to date of transition from twice-weekly HD to thrice-weekly HD.

**Geographic and socioeconomic diversity of the study population**
The geographic distribution of recruitment centers and dialysis units will confer adequate enrollment of a socioeconomically diverse patient population that will be representative of the general population treated with in-center HD in the United States. The following variables will be collected to portray participants' diversity: race, medical insurance, highest education status, employment status, annual income (at most recent employment), and residence zip code. Participants' race will be extracted through review of the electronic medical records and/or patient reports. Medical insurance will be obtained from electronic medical records and

will be categorized as Medicare, Medicaid, and/or Private, or None. Highest education status will be separated into middle school, high school, college, and university. Employment status and annual income (at most recent employment), for both patients and caregivers participants, will be collected at baseline and semi-annually. Patients' primary residence will be defined as the ZIP code at the time of enrollment and will be obtained through review of the electronic medical records and/or patient reports.
*Time Frame*: Baseline and semi-annual questions that obtain employment status.

**Rural or urban residence**
This will be determined by the use of rural-urban commuting area (RUCA) codes [1]. This method combines measures of urbanization, population density, and urban commuting to categorize locations by assigning a numeric code from 1 (most urban) to 10 (most rural). As in other studies, we will classify each patient in the current analysis as belonging to 1 of 3 mutually exclusive RUCA groups: metropolitan (RUCA, 1.0-3.9, cities with population of >50,000 and their associated suburban areas); micropolitan (RUCA, 4.0-6.0, towns or cities with population of 10,000-50,000); and rural (RUCA, >6.0, towns with population of <10,000) [2]. The ZIP code of the dialysis unit where the patient receives care will be collected to calculate the distance or travel time required between the patient's residence and dialysis unit. The *distance between a patient's residence and dialysis unit* will be calculated by measuring the direct vector distance between the geo-coded coordinates for each ZIP code, which were obtained from the Zip Code Database Project [3]. Distance between the patient's residence and dialysis unit will be classified using categories corresponding to the 0 to 50th, higher than 50th to 75th, higher than 75th to 95th, higher than 95th to 99th, and higher than 99th percentiles. Statistical analyses will proceed as described in the proposal, with appropriate terms (main effect and interaction terms) assess heterogeneity of treatment effect (HTE) by residence category and travel distance.
*Time Frame*: Baseline.

## Adverse Events (AEs) of Special Interest

The following events will be classified as serious adverse events (SAEs) of special interest and will be summarized for both treatment groups, irrespective of adjudication regarding relatedness or expectedness to study treatment:

- Hyperkalemia: pre-dialysis serum potassium (K) ≥ 6.0 mEq/L
- Hyponatremia: pre-dialysis serum sodium (Na) ≤ 125 mEq/L
- Metabolic acidosis: pre-dialysis serum total $CO_2$ ≤ 15 mEq/L
- Azotemia: pre-dialysis blood urea nitrogen (BUN) ≥ 90 mg/dL
- Anemia: pre-dialysis hemoglobin (Hb) ≤ 7.0 g/dL
- Hyperphosphatemia: pre-dialysis serum phosphorus ≥ 7.0 mg/dL
- Ultrafiltration rate: ≥ 11 mL/kg/h or ≥ 13 mL/kg/h
- Interdialytic weight gain: ≥ 4% of target weight
- Residual weight: ≥ 4% of target weight
- Pre-dialysis systolic blood pressure (SBP): ≥ 180 mmHg
- Intradialytic peak SBP: ≥ 180 mmHg
- Intradialytic nadir SBP: ≤ 90 mmHg
- Missed treatments: ≥ 30% of expected hemodialysis sessions
- Cause-specific hospitalization: related to volume overload, cardiovascular, or cerebrovascular events
- Cause-specific ED visit: related to volume overload, cardiovascular, or cerebrovascular events
- Cause-specific death: sudden death

**References**

1. Measuring rurality: rural-urban commuting area codes. Economic Research Service, United States Department of Agriculture. Available at website: ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx

2. Tonelli M, Klarenbach S, Rose C, Wiebe N, Gill J: Access to kidney transplantation among remote- and rural-dwelling patients with kidney failure in the United States. *Jama* 2009, 301(16):1681-1690.

3. The Zip Code Database Project. Available at zips.sourceforge.net

**Appendix 5. Statistical power under alternative effect sizes or non-inferiority margins**

We examined alternative combinations of the true effect size (incidence rate ratio, IRR) and non-inferiority (NI) margin, using the negative-binomial framework. These scenario analyses are intended for planning context only and do not alter the primary estimand or design assumptions (NI margin = 1.20, IRR = 0.90, one-sided α = 0.025, ≥85% power with N = 350). These scenario analyses evaluate statistical power if either (1) the true IRR difference between treatment arms is larger than originally assumed (i.e., closer to equivalence); (2) a slightly wider NI margin is used; and (3) an average annual dropout rate of 25%.

- Total N = 284 (212 effective; 106 per arm).
  If the true treatment difference is closer to equivalence (IRR = 0.90), the study retains 85% power under a modestly wider NI margin of 1.25. If the true effect is somewhat more favorable (IRR = 0.85), power is 93% under NI 1.25, and 87% even under the original NI 1.20.

  *Interpretation:* At ~80% of the original sample size, adequate power is preserved either when the true IRR is 0.90 with NI 1.25, or when the true IRR improves to 0.85 (even at NI 1.20).

- Total N = 310 (232 effective; 116 per arm).
  Under NI 1.25, power exceeds 88–95% for IRR = 0.90–0.85. With the original NI = 1.20, power is 90% for IRR = 0.85.

  *Interpretation:* With N ~310, the study remains well-powered across plausible effects; even if the true IRR is 0.90, widening NI to 1.25 restores ≥88% power.

- Sensitivity to a more permissive NI margin (NI = 1.30).
  For totals in the 284–310 range, NI = 1.30 yields high power across all IRR assumptions (e.g., 0.82–0.97 for IRR = 0.95–0.85 at N = 284; 0.85–0.98 at N = 310), indicating that a modest relaxation of the NI margin provides substantial protection against under-enrollment while maintaining rigorous inference thresholds.

| **Statistical power under alternative effect size or non-inferiority margin** | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| **Total Patient Sample Size** | **Effective sample size (annual dropout rate, 25%)** | **N per arm** | **Non-inferiority Margin = 1.20** | | | **Non-inferiority Margin = 1.25** | | | **Non-inferiority Margin = 1.30** | | |
| | | | **IRR=0.85** | **IRR=0.90** | **IRR=0.95** | **IRR=0.85** | **IRR=0.90** | **IRR=0.95** | **IRR=0.85** | **IRR=0.90** | **IRR=0.95** |
| 200 | 150 | 75 | 0.74 | 0.59 | 0.44 | 0.83 | 0.71 | 0.56 | 0.89 | 0.80 | 0.68 |
| 228 | 170 | 86 | 0.79 | 0.65 | 0.48 | 0.87 | 0.76 | 0.61 | 0.93 | 0.85 | 0.73 |
| 256 | 190 | 95 | 0.83 | 0.70 | 0.53 | 0.91 | 0.81 | 0.66 | 0.95 | 0.88 | 0.78 |
| 284 | 212 | 106 | 0.87 | 0.74 | 0.57 | 0.93 | 0.85 | 0.71 | 0.97 | 0.92 | 0.82 |
| 310 | 232 | 116 | 0.90 | 0.78 | 0.61 | 0.95 | 0.88 | 0.75 | 0.98 | 0.94 | 0.85 |
| 338 | 252 | 126 | 0.92 | 0.81 | 0.65 | 0.96 | 0.90 | 0.78 | 0.99 | 0.96 | 0.88 |
| 366 | 274 | 137 | 0.94 | 0.85 | 0.68 | 0.98 | 0.92 | 0.82 | 0.99 | 0.97 | 0.90 |

**Appendix 6: Simulation approach and estimated power to assess heterogeneity treatment effects**

We use a negative binomial model with terms for the treatment group, subgroup variable, and treatment by subgroup interaction term to make inference about heterogeneity of treatment effects. Subgroup results will be reported only when the interaction term is significant.[1] Therefore, power analysis focuses primarily on the power for the significance test for the interaction term. Power for the interaction term was determined using a simulation approach that involves the following steps:

1. Assume N=350, and generate treatment assignments using a Bernoulli distribution with probability p=0.5;

2. For each treatment group, generate subgroup membership using preliminary data from our pilot study and USRDS. The pre-specified subgroups to be explored are sex (males vs. females), age (≤65 vs. >65 years old), race (black vs. other), diabetes status (diabetic vs. not diabetic), type of vascular access (ventral venous vs. arteriovenous), and anthropometric volume (Watson volume <35L vs. ≥35L). The distribution for each subgroup is shown in all Tables (See Proportion column). Note that we do not show power estimates for age, race, and anthropometric volume because, coincidentally, they happen to have the same (55% vs. 45%) breakdown as sex. Therefore, power estimates are the same as those listed for the comparison of the treatment effect between male and female patients in the Tables.

3. We considered two types of interaction effects for completeness: quantitative and qualitative. Quantitative interactions assume a significant treatment effect, but effects are stronger in one group compared to the other. Qualitative interactions describe situations where there are no treatment effects because effects are of equal and opposite directions in the two groups.[2]

4. Simulate outcomes from a Negative Binomial distribution mean parameter given by $\mu = \beta_1 T + \beta_2 G + \beta_{12} T \times G$ for the quantitative interaction model, and by $\mu = \beta_2 G + \beta_{12} T \times G$, for the qualitative interaction model. For the quantitative model $\beta_1$was set to -0.5; $\beta_2$ was set to 0.5 for both models.

5. $\beta_{12}$ is the parameter of interest for this inference. Standardized values of $\beta_{12}$ were 0.5, 0.75, and 1. We used a dispersion parameter of 3 for all simulations. This value allowed us to maintain the interaction effect size at or near the preset value with a very tight confidence interval around it.

6. Following the data generation process, two Negative Binomial regressions are fitted: a full model (2 main effects, plus the interaction term) and a restricted model that excluded the interaction terms. We then perform a likelihood ratio test to compare the full model to the restricted model.

7. Power is determined as the number of times the null hypothesis was rejected in 1000 iterations.

8. We consider two two-sided significance levels (5% and 10%) and evaluated statistical power for uncorrected and Bonferroni corrected tests. For Bonferroni corrected significance levels, we assume six independent tests for the effects of age, sex, race, diabetes status, type of vascular access, and anthropometric volume.

For completeness, we also provide the power in each subgroup for a non-inferiority margin of 1.2.

Tables 1A and 1B below provide estimated power to detect an interaction effect at nominal significance levels of 5 and 10%, respectively. Table 1A shows that for a nominal significance level of 5%, and when the distribution of patients between the two subgroups is around 60-40% (diabetes) or 55-45% (sex), we will achieve at least 70% power to detect interactions with standardized effect sizes of 0.75 or higher whenever the interaction effect follows a quantitative model. Power improves significantly for qualitative interactions, reaching at least 89.4% when the imbalance between subgroups is not too large. However, imbalance between subgroups significantly reduces power under both models of interaction. Adjustment for multiple testing

significantly affects power estimation. However, adjustment for the correlation between the grouping variables could reduce the effective number of tests, leading to improved power.

Most clinical trials reporting subgroup analyses use higher significance levels (e.g., alpha=10, 15, or even 20%). For completeness, we provide power estimates, assuming a nominal significance threshold of 10%. Table 1B shows that the proposed sample size will provide ≥85% power to detect effect sizes of 0.75 or higher for quantitative interaction effects without adjustment for multiple testing. With a strict Bonferroni adjustment for six tests, power will drop ≥61% for comparison with a small imbalance (e.g., 55% vs. 45; 60% vs. 40%) between subgroups. Similarly to Table 1A, power drops significantly for the vascular access analysis, for which the group proportions are (80% vs. 20%).

Finally, we note that we will be powered to detect treatment effects in the stratified analyses, especially for stratification variables with little imbalance between subgroups. Table 2 shows that the proposed subgroup analyses will be well-powered to identify treatment effects across sex, age, diabetes status, anthropometric volume, etc., but we would be underpowered to identify meaningful differences among those using arteriovenous access.

In summary, the proposed sample provides adequate power to detect moderate to large differences in treatment effects across subgroups. Power is greater when the distribution of patients across the subgroups is balanced or in presence of a small imbalance (e.g., 55% vs. 45; 60% vs. 40%). We do stress that results from the proposed subgroups are meant to be exploratory. Confirmatory studies will be needed to validate them.

**Table 1A: Estimated Power for interaction term in subgroup analyses, assuming a family-wise error rate of 5%.**

| Stratification variable | Subgroup | Proportion | Sample Size | No correction for multiple testing, α=0.05 | | | | | | Bonferroni correction for 6 tests, $\alpha=8.3\times10^{-3}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Quantitative | | | Qualitative | | | Quantitative | | | Qualitative | | |
| | | | | ***E.S.=0.5*** | ***0.75*** | ***1*** | ***0.5*** | ***0.75*** | ***1*** | ***0.5*** | ***0.75*** | ***1*** | ***0.5*** | ***0.75*** | ***1*** |
| Sex | Males | 0.45 | 158 | 0.462 | 0.764 | 0.908 | 0.60 | 0.93 | 0.996 | 0.354 | 0.528 | 0.776 | 0.232 | 0.652 | 0.948 |
| | Females | 0.55 | 192 | | | | | | | | | | | | |
| Diabetes mellitus | Yes | 0.6 | 210 | 0.428 | 0.732 | 0.918 | 0.528 | 0.894 | 0.992 | 0.268 | 0.48 | 0.750 | 0.214 | 0.618 | 0.942 |
| | No | 0.4 | 140 | | | | | | | | | | | | |
| Vascular Access | Ventral Venous | 0.8 | 280 | 0.312 | 0.580 | 0.738 | 0.456 | 0.774 | 0.934 | 0.178 | 0.344 | 0.532 | 0.144 | 0.376 | 0.816 |
| | Arteriovenous | 0.2 | 70 | | | | | | | | | | | | |

**Table 1B: Estimated Power for interaction term in subgroup analyses, assuming a family-wise error rate of 10%.**

| Stratification variable | Subgroup | Proportion | Sample Size | No correction for multiple testing, α=0.10 | | | | | | Bonferroni correction for 6 tests, $\alpha=1.67x10^{-2}$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | | Quantitative | | | Qualitative | | | Quantitative | | | Qualitative | | |
| | | | | *E.S.=0.5* | *0.75* | *1* | *0.5* | *0.75* | *1* | *0.5* | *0.75* | *1* | *0.5* | *0.75* | *1* |
| Sex | Males | 0.45 | 158 | 0.57 | 0.84 | 0.97 | 0.73 | 0.95 | 0.99 | 0.30 | 0.61 | 0.87 | 0.41 | 0.81 | 0.98 |
| | Females | 0.55 | 192 | | | | | | | | | | | | |
| Diabetes mellitus | Yes | 0.6 | 210 | 0.55 | 0.83 | 0.96 | 0.72 | 0.95 | 0.99 | 0.27 | 0.60 | 0.83 | 0.42 | 0.81 | 0.98 |
| | No | 0.4 | 140 | | | | | | | | | | | | |
| Vascular Access | Ventral Venous | 0.8 | 280 | 0.43 | 0.67 | 0.86 | 0.51 | 0.85 | 0.96 | 0.16 | 0.38 | 0.64 | 0.26 | 0.62 | 0.86 |
| | Arteriovenous | 0.2 | 70 | | | | | | | | | | | | |

**Table 2: Estimated Power for stratified analyses at the 2.5% significance level**

| Stratification variable | Subgroup | Proportion | Sample Size | Stratified analyses | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Non-inferiority margin=1.2 | | | | |
| | | | | ***IRR=0.75*** | ***0.8*** | ***0.85*** | ***0.9*** | ***0.95*** |
| Sex | Males | 0.45 | 158 | 0.94 | 0.87 | 0.76 | 0.62 | 0.46 |
| | Females | 0.55 | 192 | 0.98 | 0.93 | 0.84 | 0.70 | 0.53 |
| Diabetes mellitus | Yes | 0.6 | 210 | 0.98 | 0.95 | 0.87 | 0.74 | 0.57 |
| | No | 0.4 | 140 | 0.66 | 0.54 | 0.43 | 0.32 | 0.23 |
| Vascular Access | Ventral Venous | 0.8 | 280 | 0.92 | 0.83 | 0.71 | 0.56 | 0.41 |
| | Arteriovenous | 0.2 | 70 | 0.40 | 0.32 | 0.25 | 0.19 | 0.14 |


1. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized trials: risks of subgroup-specific analyses; power and sample size for the interaction test. *J Clin Epidemiol.* 2004;57(3):229-236.
2. Wang X, Piantadosi S, Le-Rademacher J, Mandrekar SJ. Statistical Considerations for Subgroup Analyses. *Journal of Thoracic Oncology.* 2021;16(3):375-380.

## Appendix 7: Primary analysis Estimand attributes for the primary outcome and main secondary outcomes

| Outcome | Endpoint | Population | Intercurrent Events | Population Level Summary |
|---|---|---|---|---|
| **Primary:** All-cause death, ED visits, or hospitalizations | Cumulative incidence rate of all-cause death, ED visits, or hospitalizations, from the date of randomization to the end-of-study event or study closure. | All adults who meet our inclusion and exclusion criteria, which are described in Table 1 of the manuscript.* | IE1: Patient participants remain in the analysis within their originally-randomized group.<br>IE2, IE3, and IE5: Outcomes that comprise the composite primary outcomes are still monitored, and patients remain in the analysis within their originally-randomized group.<br>IE4: Patients remain in their originally-randomized group for the ITT analysis, with sensitivity analyses in the PP and AT population sets. | Incidence rate ratio (IRR), comparing the cumulative incidence rates between patients randomized to CHD, compared to those randomized to CMIHD. |
| **Main Secondary** | | | | |
| Illness Intrusiveness Rating Scale (IIRS) score | Change from baseline IIRS score. This endpoint will be calculated as the difference between the most recent IIRS score observed before an end-of-study event, and the IIRS score observed pre-randomization | All adults who meet our inclusion and exclusion criteria, which are described in Table 1 of the manuscript. | IE1: Patient participants remain in the analysis within their originally-randomized group.<br>E2 and IE3: Patient participants are right-censored at the time of the event. | Difference in the change from baseline between patients randomized to CHD, compared to those randomized to CMIHD. Change is calculated as the difference between the baseline observation and the most recent score obtained before the censoring event |
| Illness Intrusiveness Rating Scale (IIRS) score – to the point of transition or crossover | Change from baseline IIRS score. This endpoint will be calculated as the difference between the most recent IIRS score observed before transition or crossover, and the IIRS score observed pre-randomization | All adults who meet our inclusion and exclusion criteria, which are described in Table 1 of the manuscript. | IE1: Patient participants remain in the analysis within their originally-randomized group.<br>IE4: Patients remain in their originally-randomized group for the ITT analysis, with sensitivity analyses in the PP and AT population sets. | Difference in the change from baseline between patients randomized to CHD, compared to those randomized to CMIHD. Change is calculated as the difference between the baseline observation and the most recent score, with an indicator for the treatment received (AT). |
| Change in urine volume | Change from baseline urine volume. This endpoint will be calculated as the difference between the most recent urine volume observed before an end-of-study event, and the urine volume observed pre-randomization | All adults who meet our inclusion and exclusion criteria, which are described in Table 1 of the manuscript. | IE1: Patient participants remain in the analysis within their originally-randomized group.<br>E2 and IE3: Patient participants are right-censored at the time of the event. | Difference in the change from baseline between patients randomized to CHD, compared to those randomized to CMIHD. |
| Change in urine volume – up to the point of transition or crossover | Change from baseline urine volume. This endpoint will be calculated as the difference between the most recent urine volume observed beforethe point of transition or crossover, and the urine volume observed pre-randomization | All adults who meet our inclusion and exclusion criteria, which are described in Table 1 of the manuscript. | IE1: Patient participants remain in the analysis within their originally-randomized group.<br>IE4: Patients remain in their originally-randomized group for the ITT analysis, with sensitivity analyses in the PP and AT population sets. | Difference in the change from baseline between patients randomized to CHD, compared to those randomized to CMIHD. Change is calculated as the difference between the baseline observation and the most recent score, with an indicator for the treatment received (AT). |

*: Note that this definition refers to the broader patient population, not just those who will be enrolled in this trial.

**Intercurrent Definition and additional comments**

| | |
|---|---|
| IE1 | Temporary treatment outside the study team's follow-up capability for ≤6 weeks (e.g., due to travel), followed by return to treatment under the study team's monitoring. |
| IE2 | Transition to home dialysis (home HD or peritoneal dialysis) within the study team's follow-up capability, with continued limited data collection. |
| IE3 | Patient participant refusal to continue timed urine collections or questionnaire completion, with continued limited data collection. |
| IE4 | Progression from twice-weekly to thrice-weekly HD in the CMIHD group. |
| IE5 | Crossover from thrice-weekly to twice-weekly HD in the CHD group. |

### Appendix 9 -- R code for analyzing the Main Secondary Outcomes

```
rm(list=(ls()))
library(tidyverse)
library(lme4)
library(survival)
library(glmmTMB)  #For mixed-effect negative binomial
library(nortest) # For Anderson-Darling tests
library(influence.ME) # Calculate influence statistics package for mixed models

# id: patient id
# x: randomized treatment
# z: site
# access: type of vascular access
# y0: baseline value of outcome
# y: repeated measurements of outcome, with time going from 1 to k_i for patient i,
## where k_i is the number observations for patient k
# time : observation time

# Function to fit Mixed-Effect with patient follow up time as the offset

### Secondary Outcome data is loaded.

SecOutcome <-NULL

Baseline<-trial_data %>%
  filter(time==0) %>%
  mutate(baseline=outcome)%>%
  select(site,subject_id,baseline,access)

Restdat<-trial_data %>%
  filter(!time==0)

SecOutcome<-Restdat %>%
  left_join(Baseline,by=c("site","subject_id")) %>%
  select(site,subject_id,treatment,outcome,baseline,time,sex,access)

colnames(SecOutcome)<-c("z","id","x","y","y0","time","sex","access")

fit_me <- function(data) {
  tryCatch({

    model <- lmer(y ~ x + y0 + access + (1|z) + (1|id) + time, data=data, REML=FALSE,
                  control = lmerControl(
                    optimizer = "bobyqa",
                    optCtrl = list(maxfun = 100000)
                  )) ### REML initially set at FALSE for evaluating fixed effects and TRUE for random effects
```

```
    # Check model estimates
    has_na_beta <- any(is.na(fixef(model)))
    has_na_theta <- any(is.na(model@theta))
    if (has_na_beta || has_na_theta) {
      return(list(beta = NA, p_value = NA, converged = FALSE))
    }

    #Check for convergence
    conv_code <- model@optinfo$conv$opt
    if (!(conv_code == 0)){
      return(list(beta = NA, p_value = NA, converged = FALSE))
    }

    #check for singular matrix
    is_sing <- isSingular(model)

    if(is_sing){
      return(list(beta = NA, p_value = NA, converged = FALSE))
    }


    fixed_effects <- fixef(model)


    # Get p-value and beta estimate for x
    if ("x" %in% fixed_effects) {
      std_error<-sqrt(diag(vcov(model)))
      beta_est<-fixed_effects[names(fixed_effects)=="x"]
      se<-std_error[names(std_error)=="x"]
      p_value<-1-pt(abs(beta_est/se),1)
      print('i am at step 4')

      # Final check for valid results
      if(is.na(beta_est) || is.na(p_value) || !is.finite(beta_est) || !is.finite(p_value)) {
        return(list(beta = NA, p_value = NA, converged = FALSE))
      }
    } else {
      return(list(beta = beta_est, p_value = p_value, converged = TRUE))
    }
  }, error = function(e) {
    cat('found and error', e$message, "\n")
    return(list(beta = NA, p_value = NA, converged = FALSE))
  })
}


# Comprehensive diagnostic function
run_diagnostics <- function(model, data) {
```

```
cat("=== LME Diagnostic Report ===\n\n")

# 1. Normality of residuals
cat("1. OVERDISPERSION TEST\n")
residuals <- resid(model, type = "pearson")
fitted <- fitted(model)
if(length(residuals)<5000){
  shapiro_resid<-shapiro.test(residuals)
}

cat(sprintf("   Shapiro-Wilk test: W = %.4f, p = %.4f\n",
            shapiro_resid$statistic, shapiro_resid$p.value))

if (shapiro_resid$p.value > 0.05) {
  cat("Residuals appear normally distributed\n")
} else {
  cat("Residuals may deviate from normality\n")
}

# 2. Homoscedasticity
cat("\n2. HOMOSCEDASTICITY (constant variance)\n")

# Breusch-Pagan test approximation
residual_var_model <- lm(abs(residuals) ~ fitted)
bp_stat <- summary(residual_var_model)$fstatistic[1]
bp_p <- pf(bp_stat,
           summary(residual_var_model)$fstatistic[2],
           summary(residual_var_model)$fstatistic[3],
           lower.tail = FALSE)

cat(sprintf("Variance vs fitted test: F = %.3f, p = %.4f\n",
            bp_stat, bp_p))

if (bp_p > 0.05) {
  cat("No strong evidence of heteroscedasticity\n")
} else {
  cat("Possible heteroscedasticity detected\n")
}

# 3. Random effects normality
cat("3. RANDOM EFFECTS NORMALITY\n")

re_site <- ranef(model)$z[[1]]
re_subject <- ranef(model)$id[[1]]

sw_site <- shapiro.test(re_site)
cat(paste("Shapiro-Wilk (Site): p =", round(sw_site$p.value, 4), "\n"))

if (length(re_subject) <= 5000) { ## the shapiro.test() function throws an error with more than 5,000 observations.
```

```
  sw_subj <- shapiro.test(re_subject)
  cat(paste("Shapiro-Wilk (Subject): p =", round(sw_subj$p.value, 4), "\n"))
}
cat("\n")

# 4. Model fit
cat("4. MODEL FIT\n")
cat(paste("AIC:", round(AIC(model), 2), "\n"))
cat(paste("BIC:", round(BIC(model), 2), "\n"))
cat("\n")

# 5. Variance components
cat("5. VARIANCE COMPONENTS\n")
print(VarCorr(model))
cat("\n")

#5. Visual inspections
ad.test(re_site) ### formal test of normality of sites
ad.test(re_subject) ### formal test of normality of patients

## plot histograms
# Visual assessment with histograms
par(mfrow = c(1, 2))
hist(re_site, breaks = 20, main = "Site Random Effects",
     xlab = "Random Effect Value", col = "lightblue")
hist(re_subject, breaks = 30, main = "Subject Random Effects",
     xlab = "Random Effect Value", col = "lightgreen")

data$resid_pearson <- resid(model, type = "pearson")
data$resid_deviance <- resid(model, type = "deviance")
data$fitted_values <- fitted(model)

# Plot Pearson residuals vs fitted values
ggplot(data, aes(x = fitted_values, y = resid_pearson)) +
  geom_point(alpha = 0.5) +
  geom_hline(yintercept = 0, color = "red", linetype = "dashed") +
  geom_smooth(se = TRUE, color = "blue") +
  labs(title = "Pearson Residuals vs Fitted Values",
       x = "Fitted Values",
       y = "Pearson Residuals") +
  theme_minimal()

# Plot deviance residuals
ggplot(data, aes(x = fitted_values, y = resid_deviance)) +
  geom_point(alpha = 0.5) +
  geom_hline(yintercept = 0, color = "red", linetype = "dashed") +
  geom_smooth(se = TRUE, color = "blue") +
  labs(title = "Deviance Residuals vs Fitted Values",
       x = "Fitted Values",
```

```
         y = "Deviance Residuals") +
    theme_minimal()

  #6. Influence Measures

  # Calculate Cook's distance for each observation
  infl <- influence(model, obs = TRUE)
  data$cooks_d <- cooks.distance(infl)

  # Plot Cook's distance
  ggplot(data, aes(x = seq_along(cooks_d), y = cooks_d)) +
    geom_segment(aes(xend = seq_along(cooks_d), yend = 0), linewidth = 0.5) +
    geom_hline(yintercept = 4/N, color = "red", linetype = "dashed") + # Threshold: 4/n
    labs(title = "Cook's Distance", x = "Observation Index", y = "Cook's Distance") +
    theme_minimal()

  # Identify influential points
  influential <- which(data$cooks_d > 4/N)
  print(paste("Number of influential observations:", length(influential)))

  # Examine influential observations -- if any.
  data[influential, ]

  # Return invisibly for further use
  invisible(list(
    dispersion = dispersion,
    shapiro_site = sw_site,
    aic = AIC(model),
    bic = BIC(model)
  ))
}

# Quick convergence check
quick_check <- function(model, model_name = "Model") {

  has_na <- any(is.na(fixef(model))) || any(is.na(model@theta))
  conv_code <- model@optinfo$conv$opt
  is_sing <- isSingular(model)

  grad <- model@optinfo$derivs$gradient
  max_grad <- if (!is.null(grad)) max(abs(grad)) else NA

  cat(sprintf("%s:\n", model_name))
  cat(sprintf("  NA params: %s, Conv: %d, Singular: %s, Grad: %.6f\n",
              has_na, conv_code, is_sing, max_grad))

  ok <- !has_na && conv_code == 0 && !is_sing &&
    (is.na(max_grad) || max_grad < 0.01)
```

```
  if (ok) {
    cat("  Status: OK for inference\n")
  } else {
    cat("  Status: Problematic\n")
  }

  return(ok)
}


# Function to perform and report LRT for fixed effects
perform_lrt_fixed <- function(model_reduced, model_full,
                              test_name = "Fixed Effect",
                              alpha = 0.05) {

  cat(sprintf("\n--- LRT: %s ---\n\n", test_name))

  # Pre-test validation
  ok_full <- quick_check(model_full, "Full model")
  ok_reduced <- quick_check(model_reduced, "Reduced model")

  reliable <- ok_full && ok_reduced

  if (!reliable) {
    cat("\nWARNING: At least one model has convergence issues\n")
    cat("Results should be interpreted with caution\n\n")
  }

  # Perform LRT
  lrt <- anova(model_reduced, model_full)
  print(lrt)

  # Extract statistics
  chi_sq <- lrt$Chisq[2]
  df <- lrt$npar[2] - lrt$npar[1]
  p_value <- lrt$`Pr(>Chisq)`[2]

  cat(sprintf("\nTest statistic: χ² = %.3f\n", chi_sq))
  cat(sprintf("Degrees of freedom: %d\n", df))
  cat(sprintf("P-value: %.4f\n", p_value))

  # Decision
  if (p_value < alpha) {
    cat(sprintf("\nDecision: Reject H₀ (p < %.2f) ✓\n", alpha))
    cat("Conclusion: Effect is statistically significant\n")
  } else {
    cat(sprintf("\nDecision: Fail to reject H₀ (p ≥ %.2f)\n", alpha))
    cat("Conclusion: Effect is not statistically significant\n")
  }
```

```
  return(list(
    chi_sq = chi_sq,
    df = df,
    p_value = p_value,
    reliable = reliable
  ))
}

################################################################################################


# Function to perform LRT for variance components
perform_lrt_variance <- function(model_reduced, model_full,
                                 test_name = "Variance Component",
                                 alpha = 0.05) {

  cat(sprintf("\n--- LRT: %s ---\n\n", test_name))

  # Pre-test validation
  ok_full <- quick_check(model_full, "Full model")
  ok_reduced <- quick_check(model_reduced, "Reduced model")

  reliable <- ok_full && ok_reduced

  if (!reliable) {
    cat("\nWARNING: At least one model has convergence issues\n")
    cat("Results should be interpreted with caution\n\n")
  }

  # Perform LRT
  lrt <- anova(model_reduced, model_full)
  print(lrt)

  # Extract statistics
  chi_sq <- lrt$Chisq[2]
  df <- lrt$npar[2] - lrt$npar[1]
  p_value_raw <- lrt$`Pr(>Chisq)`[2]

  # BOUNDARY CORRECTION: Halve p-value for variance component tests
  p_value_corrected <- p_value_raw / 2

  cat(sprintf("\nTest statistic: χ² = %.3f\n", chi_sq))
  cat(sprintf("Degrees of freedom: %d\n", df))
  cat(sprintf("Raw p-value: %.4f\n", p_value_raw))
  cat(sprintf("Corrected p-value (boundary): %.4f\n", p_value_corrected))
  cat("\nNote: P-value halved due to boundary constraint (variance ≥ 0)\n")

  # Decision using corrected p-value
```

```
  if (p_value_corrected < alpha) {
    cat(sprintf("\nDecision: Reject H₀ (p < %.2f) ✓\n", alpha))
    cat("Conclusion: Variance component is statistically significant\n")

    # Report variance
    vc <- VarCorr(model_full)
    var_names <- names(vc)
    if (length(var_names) > 0) {
      for (vn in var_names) {
        var_val <- as.numeric(vc[[vn]])
        cat(sprintf("  %s variance: %.4f\n", vn, var_val))
      }
    }
  } else {
    cat(sprintf("\nDecision: Fail to reject H₀ (p ≥ %.2f)\n", alpha))
    cat("Conclusion: Variance component not significantly different from zero\n")
  }

  return(list(
    chi_sq = chi_sq,
    df = df,
    p_value_raw = p_value_raw,
    p_value_corrected = p_value_corrected,
    reliable = reliable
  ))
}

model_full <- lmer(y ~ x + y0 + access + (1|z) + (1|id) + time + sex + sex*x, data=data, REML=FALSE,
               control = lmerControl(
                 optimizer = "bobyqa",
                 optCtrl = list(maxfun = 100000)
               )) ### REML initially set at FALSE for evaluating fixed effects and TRUE for random effects
model_restricted <- lmer(y ~ x + y0 + access + (1|z) + (1|id) + time + sex, data=data, REML=FALSE,
                   control = lmerControl(
                     optimizer = "bobyqa",
                     optCtrl = list(maxfun = 100000)
                   )) ### REML initially set at FALSE for evaluating fixed effects and TRUE for random effects

# Test 1: Treatment × Sex interaction
lrt_interaction <- perform_lrt_fixed(
  model_restricted,
  model_full,
  test_name = "Treatment × Sex Interaction"
)


model_full_subject <- lmer(y ~ x + y0 + access + (1|z) + (1|id) + time, data=data, REML=TRUE,
                   control = lmerControl(
                     optimizer = "bobyqa",
```

```
                    optCtrl = list(maxfun = 100000)
                  )) ### REML initially set at FALSE for evaluating fixed effects and TRUE for random effects
model_restricted_no_subj <- lmer(y ~ x + y0 + access + (1|z), data=data, REML=TRUE,
                          control = lmerControl(
                            optimizer = "bobyqa",
                            optCtrl = list(maxfun = 100000)
                          )) ### REML initially set at FALSE for evaluating fixed effects and TRUE for random effects

lrt_subject <- perform_lrt_variance(
  model_full_subject,
  model_restricted_no_subj,
  test_name = "Subject Random Effect"
)
```

Table 1: Demographic characteristics by treatment assignment

| Baseline Characteristic[1] | N | Overall | CMIHD | CHD |
|---|---|---|---|---|
| Sex | | | | |
| Male | | | | |
| Female | | | | |
| Age | | | | |
| Height | | | | |
| Weight | | | | |
| BSA ($m^2$) | | | | |
| BMI ($m^2$) | | | | |
| Race and ethnicity | | | | |
| Asian | | | | |
| Black or African American | | | | |
| Hispanic of any race | | | | |
| Native Hawaiian or other Pacific Islander | | | | |
| White | | | | |
| Other | | | | |
| Unknown | | | | |
| Highest education level | | | | |
| None | | | | |
| Elementary school | | | | |
| Middle school | | | | |
| High school | | | | |
| Undergraduate/College | | | | |
| Postgraduate/University | | | | |
| Smoking status | | | | |
| Current | | | | |
| Former | | | | |
| Never | | | | |
| Unknown | | | | |
| Depression diagnosis | | | | |
| No | | | | |
| Other | | | | |
| Yes | | | | |
| Anxiety diagnosis | | | | |
| No | | | | |
| Yes | | | | |

[1] Values are either mean (SD) or N (%)

Table 2: Socioeconomic characteristics and comorbidities by treatment assignment

| Socioeconomic Characteristic | N | Overall | CMIHD | CHD |
|---|---|---|---|---|
| Medical coverage | | | | |
| Medicaid | | | | |
| Medicare | | | | |
| Employer Group Health Insurance | | | | |
| Veterans Affairs | | | | |
| Medicare Advantage | | | | |
| Other | | | | |
| None | | | | |
| Unknown | | | | |
| Employment status, 6 months prior to chronic hemodialysis initiation | | | | |
| Unemployed | | | | |
| Employed full time | | | | |
| Employed part time | | | | |
| Homemaker | | | | |
| Retired (age/preference) | | | | |
| Retired (disability) | | | | |
| Medical leave of absence | | | | |
| Student | | | | |
| Unknown | | | | |
| Employment status, at chronic hemodialysis initiation | | | | |
| Unemployed | | | | |
| Employed full time | | | | |
| Employed part time | | | | |
| Homemaker | | | | |
| Retired (age/preference) | | | | |
| Retired (disability) | | | | |
| Medical leave of absence | | | | |
| Student | | | | |
| Unknown | | | | |
| Primary cause of renal failure | | | | |
| ICD codes | | | | |
| Unknown | | | | |
| Comorbid conditions | | | | |
| Congestive heart failure | | | | |
| Atherosclerotic heart disease | | | | |
| Other cardiac disease | | | | |
| Cerebrovascular disease | | | | |
| Peripheral vascular disease | | | | |
| History of hypertension | | | | |
| Diabetes | | | | |
| Chronic obstructive pulmonary disease | | | | |
| Tobacco use | | | | |
| Malignant neoplasm | | | | |
| Institutionalized | | | | |
| Assisted living | | | | |
| Nursing home | | | | |
| Other institution | | | | |

[1] Values are either mean (SD) or N (%)

Table 3: Adverse and serious adverse event summary

| | Overall | CMIHD | CHD |
|---|---|---|---|
| Days accrued since randomization [1] | | | |
| Adverse events | | | |
| Patients with $\geq$ one adverse event | | | |
| Serious adverse events | | | |
| Patients with $\geq$ one serious adverse event | | | |

[1]Values are either mean (SD) or N unless otherwise stated

Table 4: Primary and secondary outcomes by treatment assignment

| Outcome | Overall | | | CMIHD | | | CHD | | |
|---|---|---|---|---|---|---|---|---|---|
| | Patients (N) | Events | Estimate (97.5% CI) | Patients (N) | Events | Estimate (97.5% CI) | Patients (N) | Events | Estimate (97.5% CI) |
| **Follow-up duration (months)** | | | | | | | | | |
| **Primary outcomes** | | | | | | | | | |
| Rate of all-cause ED visits, hospitalizations, or death | | | | | | | | | |
| All-cause ED visit rate | | | | | | | | | |
| Hospitalization rate | | | | | | | | | |
| Death rate | | | | | | | | | |
| **Secondary outcomes** | | | | | | | | | |
| Health-related quality of life | | | | | | | | | |
| IIRS, patient/month | | | | | | | | | |
| IIRS > threshold | | | | | | | | | |
| EQ-5D-5L score, patient/semester | | | | | | | | | |
| Employment status, patient/semester | | | | | | | | | |
| Residual kidney function, patient/quarter | | | | | | | | | |
| Change in urine output (mL/24 h) | | | | | | | | | |
| Change in KUC (mL/min/1.73 $m^2$) | | | | | | | | | |
| Change in kidney creatinine clearance (mL/min/1.73 $m^2$) | | | | | | | | | |

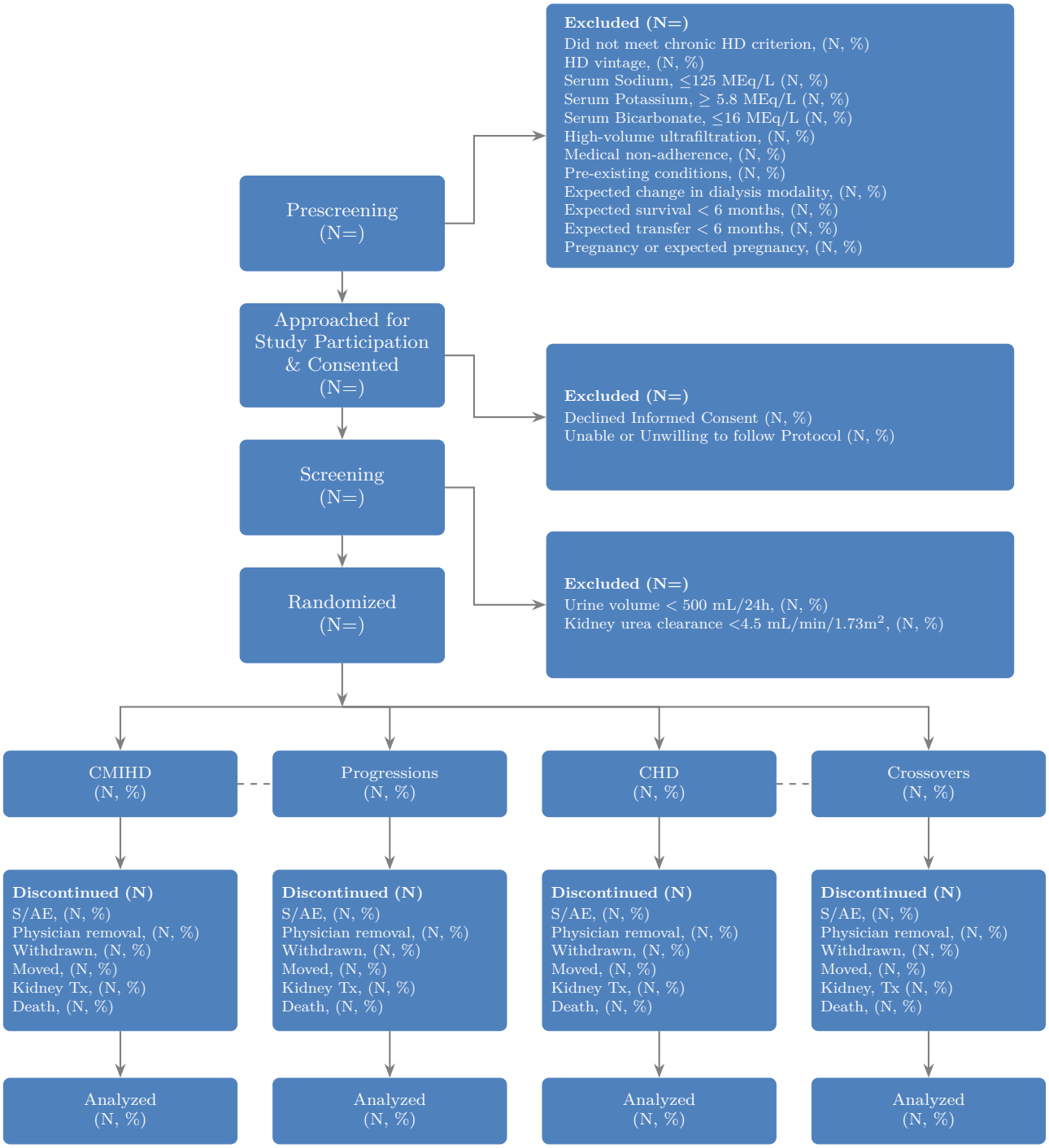


Figure 1. Consort diagram

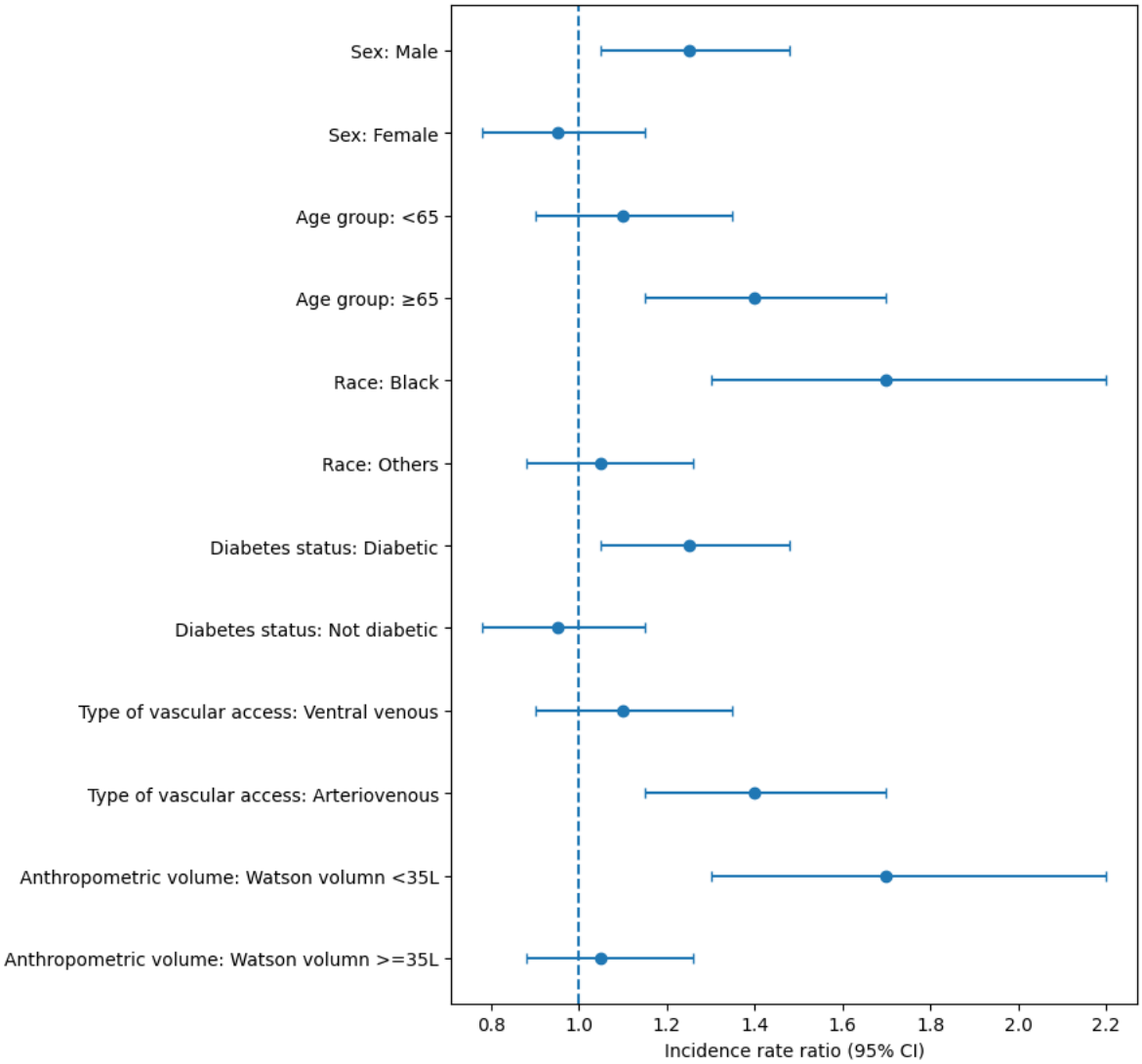


Figure 2. Primary outcome forest plot