Simple compared to covariate-constrained randomization methods in balancing baseline characteristics: a case study of randomly allocating 72 hemodialysis centers in a cluster trial

Ahmed A Al-Jaishi (Ahmed.AlJaishi@lhsc.on.ca)  
McMaster University  https://orcid.org/0000-0003-0376-2214

Stephanie N Dixon  
ICES

Eric McArthur  
ICES

PJ Devereaux  
McMaster University Department of Health Research Methods Evidence and Impact

Lehana Thabane  
McMaster University Department of Health Research Methods Evidence and Impact

Amit X Garg  
Lawson Research Institute: London Health Sciences Centre

Research Article

Keywords: cluster randomized trial, covariate-constrained, randomization, balanced allocation, restricted randomization

DOI: https://doi.org/10.21203/rs.3.rs-278523/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License
Abstract

**Background and aim:** Some parallel-group cluster randomized trials use covariate-constrained rather than simple randomization. This is done to increase the chance of balancing the groups on cluster- and patient-level baseline characteristics. This study assessed how well two covariate-constrained randomization methods balanced baseline characteristics compared with simple randomization.

**Methods:** We conducted a mock three-year cluster randomized trial, with no active intervention, that started April 1st, 2014, and ended March 31st, 2017. We included a total of 11,832 patients from 72 hemodialysis centers (clusters) in Ontario, Canada. We randomly allocated the 72 clusters into two groups in a 1:1 ratio on a single date using individual- and cluster-level data available up to April 1st, 2013. Initially, we generated 1,000 allocation schemes using simple randomization. Then, as an alternative, we performed covariate-constrained randomization based on historical data from these centers. In one analysis, we restricted on a set of 11 individual-level prognostic variables; in the other, we restricted on principal components generated using 29 baseline historical variables.

We created 300,000 different allocations for the covariate-constrained randomizations, and we restricted our analysis to the 30,000 best allocations. We then randomly sampled 1,000 schemes from the 30,000 best allocations. We summarized our results with each randomization approach as the median (25th, 75th percentile) number of balanced baseline characteristics. There were 156 baseline characteristics, and a variable was balanced when the between-group standardized difference was ≤ 10%.

**Results:** The three randomization techniques had at least 125 of 156 balanced baseline characteristics in 90% of sampled allocations. The median number of balanced baseline characteristics using simple randomization was 147 (142, 150). The corresponding value for covariate-constrained randomization using 11 prognostic characteristics was 149 (146, 151), while for principal components, the value was 150 (147, 151). The median number of balanced baseline characteristics using the two covariate-constrained randomizations were statistically different from simple randomization (p-value < 0.0001).

**Conclusion:** In this setting with 72 clusters, constraining the randomization using historical information achieved better balance on baseline characteristics compared with simple randomization; however, the magnitude of benefit was modest.

1. Introduction

The cluster randomized trial (CRT) study design is useful when the interventions are naturally implemented on groups of individuals.\(^1,2\) In contrast to individually randomized trials, CRTs randomly allocate groups rather than independent individuals. Simple randomization is the most basic and straightforward type of random allocation. Each "randomized unit" is assigned purely by chance. However, suppose the total number of randomized units is small (e.g., fewer than 20 units). In that case, simple randomization may result in a moderate to high probability of imbalance in baseline characteristics between the trial arms.\(^3\) In two-group, parallel-arm, individual-level trials, some have
suggested that including at least 1,000 participants per group is required to provide sufficient protection against the imbalance of baseline characteristics.\(^4\) In the CRT setting, it is often impossible to have such a large number of randomized units. In a systematic review of 300 CRTs, 50% of trials randomly allocated fewer than 21 clusters, and 75% allocated fewer than 52 clusters.\(^5\)

Observing between-group differences in a trial's baseline characteristics complicates the interpretation of observed treatment effects and threatens the trial's internal validity.\(^6\)–\(^8\) Other randomization techniques may help minimize the risk of imbalance on baseline measured characteristics when using parallel arm CRT designs.\(^8\) These techniques are described as "restricted" or "constrained" and include stratification, matching, minimization, and covariate-constrained randomization. All restricted methods require \textit{a priori} knowledge about participating clusters and the baseline measures used for the restriction process.

Covariate-constrained randomization can provide a better baseline balance than other allocation methods (e.g., simple random allocation, stratification, and minimization).\(^3\)\(^,\)\(^8\)\(^–\)\(^10\) This manuscript focuses on covariate-constrained randomization, where we constrained the randomization process using two sets of baseline characteristics (either constraining on a set of prognostic variables or principal components.) Principal components are a small set of artificial variables that explain most of the variance about a larger group of variables.

Covariate-constrained randomization limits the potential schemes available for selection among all possible allocations (called the randomization space). This method simultaneously balances several measured cluster- or individual-level characteristics to ensure that the two treatment arms are similar at baseline.\(^8\),\(^9\) Briefly, the covariate-constrained randomization process includes \(\text{(i)}\) \text{a priori} identifying and specifying a limited number of key prognostic cluster- or individual-level variables associated with the outcome that will be used to constrain the randomization process (or a function of baseline characteristics, for example, principal components); \(\text{(ii)}\) when there are 20 or more clusters, either enumerating all or generating at least 100,000 allocation schemes; \(\text{(iii)}\) for each allocation scheme, estimating balance on the selected baseline characteristics according to some predefined balance metric (e.g., absolute differences, standardized differences, or another measure\(^11\)); \(\text{(iv)}\) choosing a constrained randomization space containing a subset of allocations that are balanced on the constrained baseline characteristics (e.g., 10% of the best allocations\(^11\)–\(^13\)); and \(\text{(v)}\) randomly selecting one allocation scheme from the constrained randomization space that will be used for the trial.

There is a trade-off between the potential for a better balance achieved on the constrained baseline characteristics and the potential concerns with highly restricted randomization.\(^9\),\(^12\) These trade-offs can include \(\text{(i)}\) jeopardizing the appearance of impartiality, for example, if pairs of clusters always (or never) appear in the same arm; \(\text{(ii)}\) a departure from the nominal Type I error when clusters with correlated outcomes have a very high or very low probability of being included in the same trial arm; and \(\text{(iii)}\) a loss in statistical power when variables used in the constrained randomization do not associate with the trial outcome.\(^9\),\(^12\) Also, covariate-constrained randomization uses historical data on recruited clusters to
capture baseline information on demographics, patients' medical histories, and historical rates of the outcomes. However, historical data are usually several months to years old at randomization. In an "open cohort" setting, information available at the randomization date also cannot account for new participants entering the cohort during the trial period. Thus, the balance achieved at the time of randomization with historical information does not guarantee a balance of the baseline characteristics during the trial period. It is important to note that the randomization design (i.e., constrained variables) needs to be considered at the analysis stage.

We conducted this study to understand the best practices for randomizing hemodialysis centers into two parallel groups in Ontario, Canada. The lessons learned from this study will help our group make informed decisions about randomization processes for several CRTs that we plan to advance.

1.1 Motivating example

In the hemodialysis setting, the CRT is an attractive design when implementing interventions at the dialysis center level. The CRT design offers logistical and administrative advantages such as simplifying the trial organization when evaluating policy- or cluster-level intervention.

Suppose that we wish to undertake a CRT with hemodialysis centers in Ontario, Canada. In this example, we used historical data from administrative data sources to conduct covariate-constrained randomization. The trial period was three years, from April 1st, 2014, to March 31st, 2017. The primary outcome was a composite of time-to-first event for cardiovascular-related death or non-fatal major cardiovascular-related hospitalization (hospital admission for myocardial infarction, stroke, or congestive heart failure).

1.2 Objectives

This paper compared randomization methods for a two-arm, parallel-group CRT, with the intent that all individuals within a given randomized center will receive the same intervention. We randomized a moderate number of clusters (i.e., hemodialysis centers) using either simple randomization or covariate-constrained randomization with pre-trial historical records (called the Population for Randomization). We performed the randomization on a single date and allowed patients to enter the cohort throughout the study period. We compared simple randomization to covariate-constrained randomization on balance achieved on a set of baseline characteristics during a three-year trial period (called the Trial Population). We constrained either on prognostic variables or principal components.

Our secondary aim was to assess whether, in the absence of any intervention, the allocation schemes selected through the constrained randomization process preserved (i) a null treatment effect on the primary outcome and (ii) a 5% nominal Type I error rate.

2. Methods
2.1 Design and setting

We used a CRT design of outpatient hemodialysis centers in Ontario, Canada, that cared for a minimum of 15 patients.

In 2013, Ontario had approximately 13.5 million residents with universal access to healthcare and physician services. In the same period, Ontario had 26 regional dialysis programs that oversaw over 100 hemodialysis centers caring for about 8,000 in-center patients in the outpatient setting.

2.2 Data sources

We ascertained center- and patient-level characteristics using records from linked healthcare databases in Ontario, Canada (Appendix 1). These datasets were linked using unique encoded identifiers and analyzed at ICES.

2.3 Patients

We included two populations of patients, the Population for Randomization and the Trial Population. The Population for Randomization included patients who were actively receiving in-center hemodialysis on April 1st, 2013. The Trial Population included an open cohort of patients who received in-center hemodialysis on April 1st, 2014 or began receiving in-center hemodialysis during the trial period.

2.4 Baseline characteristics

We identified two cluster- and 86 individual-level (total 88) baseline characteristics to describe each cohort (Appendix 2); the cluster-level characteristics included the center size and historical rate for the primary outcome. There were 23 continuous, 58 binary, and 14 categorical baseline characteristics. Nine continuous baseline characteristics were also featured as categorical variables. We created a new binary (or "dummy") variable to indicate each level of a category's presence or absence. In total, we evaluated 156 continuous or binary candidate baseline characteristics.

2.5 Randomization process

2.5.1 Sequence generation:

We randomly allocated the 72 hemodialysis centers into two groups in a 1:1 ratio on a single date. Initially, we generated 1,000 random allocation schemes using simple (unconstrained) randomization that required no information on baseline characteristics. This number of random allocations produced an estimate within 0.5% accuracy of the true hazard ratio of 1.00 with a 5% significance level and a standard deviation of 0.08. Then, as an alternative, we performed the covariate-constrained randomization using pre-trial historical records, which ended April 1st, 2013 (see details below). Using PROC PLAN in SAS version 9.4 (SAS Institute Inc., NC Cary), we generated 300,000 unique allocation schemes of the 72 centers (Appendix 3). Greene (2017) suggested performing at least 100,000 allocations when there are at least 20 clusters; with our computational capacity, we enumerated 300,000 allocations.
2.5.2 Covariate-constrained randomization:

We performed the covariate-constrained randomization in the following series of steps using baseline characteristics of the *Population for Randomization*.\textsuperscript{5,8,9,41}

**Step 1**

Randomly selected 300,000 allocation schemes from the $4.43 \times 10^{20}$ possible allocation schemes.

**Step 2**

For each of the 300,000 allocation schemes, we restricted the randomization space using one of two constraining criteria.\textsuperscript{8}

i. We constrained the allocation on a set of 11 baseline characteristics deemed prognostic for the outcome, based on prior literature or clinical experience (Appendix 4a).

ii. We constrained the allocation on principal components. A principal component analysis is a dimensionality reduction technique whereby a dataset with many variables is transformed into a smaller set of artificial variables (called principal components). These principal components ideally retain some or most of the meaningful properties of the original set of variables. We used the principal components to account for some of the variation in the observed data and as criterion variables in our constrained randomization process (Appendix 4b).

We compared baseline differences between the two arms using standardized differences,\textsuperscript{42,43} which describes the differences between group means or proportions relative to the pooled standard deviation.

**Step 3**

For each allocation scheme from the *Population for Randomization*, we counted the number of constrained variables with a standardized difference greater than 10\% and calculated the sum of the constrained variables’ standardized differences.\textsuperscript{42,44} We added a penalty of ten units to the sum of standardized differences for each imbalanced constrained variable. We imposed this penalty to favor allocation schemes that had the least number of imbalanced constrained baseline characteristics. For example, if the sum of standardized differences was two and three constrained variables were imbalanced, the penalized sum of standardized differences would be 32.

From the 300,000 randomization schemes, we constrained the randomization space to the 30,000 best allocation schemes, based on the smallest sum of the penalized standardized differences.\textsuperscript{11–13} From the 30,000 best allocations, we randomly sampled 1,000 allocations to reduce the computational time for analysis.\textsuperscript{11,12}

2.6. Statistical Analysis
For the 1,000 sampled schemes, we (i) estimated the percentage of times each center was allocated to each arm, (ii) estimated the percentage of times each combination of center pairs appeared in the same group, and (iii) calculated the standardized difference of all 156 baseline characteristics for the Trial Population. We then estimated the percentage of time each of the 156 baseline characteristics was balanced among the 1,000 sampled randomization schemes, (iv) calculated the median (25th, 75th percentile) number of baseline characteristics balanced for the Trial Population, (v) we used the Kruskal-Wallis H Test to evaluate the null hypothesis that the population medians from the three randomization techniques are equal, and finally (vi) estimated the unadjusted and adjusted hazard ratio between the randomized arms, for the time-to-first event of the composite outcome of cardiovascular-related death or a non-fatal cardiovascular-related hospitalization (see definition of outcome in Appendix 5; this is a primary outcome for future trials that is highly relevant to patients and their providers). We estimated the hazard ratio using a generalized-estimating-equations extension for the Cox proportional hazard model, with an exchangeable covariance matrix to account for within-center clustering. For each of the 1,000 sampled randomization schemes, the models were fitted to patient-level data from the Trial Population. We conducted unadjusted and another analysis adjusting for the randomization design (i.e., adjusted analyses using the constrained baseline characteristics by adding these variables into the model).

We stopped following patients on March 31st, 2017, or earlier if they died. We summarized the hazard ratios as the mean with the 2.5th and 97.5th percentiles, which corresponded to the hazard ratio estimate with a 95% confidence interval. We expected to observe no between-group differences in the event rate of our primary outcome approximately 95% of the time (i.e., a nominal Type I error of 5%). The use of 1,000 randomizations allowed us to detect a Type I error between 3.6% and 6.4% as not significantly different than 5%; we used a standard test based on the normal approximation to the binomial distribution as described by Rosner (1995).

3. Results

3.1 Characteristics of cohorts

The Population for Randomization (n = 5,812) included all patients receiving in-center hemodialysis on April 1st, 2013. The Trial Population (n = 11,832) included patients receiving hemodialysis on April 1st, 2014 (n = 5,410) and patients who started in-center hemodialysis during the three-year trial period (n = 6,412). The Trial Population included 4,415 patients (37%) that were also in the Population for Randomization. The median (25th, 75th percentile) number of patients in each center for the Population for Randomization was 61 (28, 105) and for the Trial Population was 131 (55, 227).

The Population for Randomization and the Trial Population differed on several baseline characteristics (Table 1 and Appendix 2). However, the differences were mostly attributed to the inherent differences between prevalent and new patients starting hemodialysis (e.g., length of time on dialysis, number of dialysis sessions in the prior year, healthcare services utilization, and general practitioner visits the preceding year.)
Table 1
Select baseline characteristics. The Population for Randomization included patients on hemodialysis as of April 1, 2013. The Trial Population included an open cohort of patients receiving in-center hemodialysis on April 1, 2014, or began receiving in-center hemodialysis during the trial period between April 1, 2014, and March 31, 2017.

| Baseline characteristic                                | VALUE                                                                 | Population for Randomization | Trial Population  |
|--------------------------------------------------------|----------------------------------------------------------------------|------------------------------|-------------------|
| Centers                                                | Number of centers (n Patients)                                       | 72 (n = 5,812)               | 72 (n = 11,832)   |
| Center Size                                            | Mean ± Standard deviation                                            | 81 (69)                      | 164 (137)         |
| The composite outcome of CV-related death or major CV-related hospitalization | Historic rate per 100 person-years (cluster standard deviation)       | 10 (3.7)                     | 11 (3.3)          |
| Age (years)                                            | Mean ± Standard deviation                                            | 67 (15)                      | 66 (15)           |
| Sex                                                    | Male                                                                 | 3373 (58%)                   | 7069 (60%)        |
| Living in a rural area                                 | Yes                                                                  | 359 (6%)                     | 809 (7%)          |
| Etiology for End-stage kidney disease                  | Diabetes                                                             | 2194 (38%)                   | 4472 (38%)        |
|                                                       | Glomerulonephritis/autoimmune diseases                               | 882 (15%)                    | 1575 (13%)        |
|                                                       | Drug-induced nephropathy                                             | 83 (1%)                      | 159 (1%)          |
|                                                       | Polycystic kidney disease                                            | 229 (4%)                     | 426 (4%)          |
|                                                       | Renal vascular disease                                               | 1115 (19%)                   | 1852 (16%)        |
|                                                       | Other                                                                | 738 (13%)                    | 2221 (19%)        |
|                                                       | Unknown                                                              | 571 (10%)                    | 1127 (10%)        |
| Race                                                   | Asian                                                                | 475 (8%)                     | 854 (7%)          |
|                                                       | Black                                                                | 562 (10%)                    | 1022 (9%)         |
|                                                       | White                                                                | 3698 (64%)                   | 7598 (64%)        |
|                                                       | Other                                                                | 1038 (18%)                   | 2173 (18%)        |
|                                                       | Unknown                                                              | 39 (1%)                      | 185 (2%)          |
| First dialysis modality                               | Home hemodialysis                                                    | 33 (1%)                      | 84 (1%)           |
| Baseline characteristic                        | VALUE                        | Population for Randomization | Trial Population |
|-----------------------------------------------|------------------------------|------------------------------|------------------|
| In-center hemodialysis                        | 5215 (90%)                   | 10529 (89%)                  |
| Peritoneal dialysis                           | 564 (10%)                    | 1219 (10%)                   |
| First type of vascular access                 |                              |                              |
| Arteriovenous graft                           | 103 (2%)                     | 198 (2%)                     |
| Arteriovenous fistula                         | 1044 (18%)                   | 1999 (17%)                   |
| Central venous catheter                       | 3927 (68%)                   | 8157 (69%)                   |
| Peritoneal catheter                           | 456 (8%)                     | 1092 (9%)                    |
| Unknown                                       | 282 (5%)                     | 386 (3%)                     |
| Last vascular access used before the index date|                              |                              |
| Arteriovenous graft                           | 223 (4%)                     | 376 (3%)                     |
| Arteriovenous fistula                         | 2159 (37%)                   | 3461 (29%)                   |
| Central venous catheter                       | 3376 (58%)                   | 7544 (64%)                   |
| Patient in the Ontario Drug Benefit in the six months before the index date | Yes ⁴ | 4494 (86%) | 10196 (86%) |
| Number of unique hypertensive prescriptions in the six months before the index date | Mean ± Standard deviation | 2 (2) | 2 (2) |
| Prescribed hypertensive drugs ³               |                              |                              |
| Angiotensin-converting enzyme (ACE) inhibitors| 1157 (26%)                   | 2199 (22%)                   |
| Angiotensin II Receptor Blocker               | 1258 (28%)                   | 2413 (24%)                   |
| Beta-Blockers                                 | 2649 (59%)                   | 5604 (55%)                   |
| Calcium Channel Blocker                       | 2392 (53%)                   | 5453 (53%)                   |
| Diuretics                                     | 1612 (36%)                   | 4242 (42%)                   |
| Prior CABG/PCI                                | Yes                          | 1234 (21%) | 2612 (22%) |
| Coronary Artery Disease (with angina)         | Yes                          | 3541 (61%) | 6861 (58%) |
| Baseline characteristic                        | VALUE          | Population for Randomization | Trial Population |
|-----------------------------------------------|----------------|-------------------------------|------------------|
| Heart failure                                 | Yes            | 2862 (49%)                    | 6177 (52%)       |
| Diabetes mellitus                             | Yes            | 3402 (59%)                    | 7244 (61%)       |
| Depression                                    | Yes            | 1528 (26%)                    | 3161 (27%)       |
| Ischemic Stroke                               | Yes            | 261 (4%)                      | 551 (5%)         |
| Lower extremity amputation                    | Yes            | 298 (5%)                      | 554 (5%)         |
| Lung disease (COPD)                           | Yes            | 2276 (39%)                    | 4666 (39%)       |
| Myocardial infarction                         | Yes            | 1454 (25%)                    | 2909 (25%)       |
| Major Cancer                                  | Yes            | 841 (14%)                     | 1803 (15%)       |
| Peripheral vascular disease                   | Yes            | 1699 (29%)                    | 3055 (26%)       |
| Modified Charlson comorbidity score \(^5\)    | Mean ± Standard deviation | 4 (2)                        | 4 (2)            |
|                                              | 2              | 1931 (33%)                    | 3631 (31%)       |
|                                              | 3              | 519 (9%)                      | 1028 (9%)        |
|                                              | 4              | 1382 (24%)                    | 2883 (24%)       |
|                                              | 5+             | 1980 (34%)                    | 4290 (36%)       |
| Having a kidney transplant before the index date | Yes            | 71 (1%)                       | 148 (1%)         |
| Number of days spent in long-term care in the year before the index date | Mean ± Standard deviation | 37 (164)                     | 25 (133)         |
| Number of days spent in the hospital in the year before the index date | Mean ± Standard deviation | 11 (28)                      | 17 (31)          |
| Number of hospital admissions in the year before the index date | Mean ± Standard deviation | 1 (1)                        | 1 (1)            |
| Time since starting dialysis (days)           | Mean ± Standard deviation | 1847 (1836)                  | 1327 (1782)      |
Baseline characteristic | VALUE | Population for Randomization | Trial Population
--- | --- | --- | ---
1 Population for Randomization included patients that were on hemodialysis as of April 1, 2013 index date. The Trial Population included patients on hemodialysis as of April 1, 2014, and any patient who started in-center hemodialysis at one of the 72 participating centers during the three-year trial period. Follow-up ended March 31, 2017. The index date was the first date patients entered the respective cohort.

2 The composite outcome of cardiovascular-related death or hospitalization for myocardial infarction, ischemic stroke, congestive heart failure.

3 Percentages presented only for patients eligible to receive the Ontario Drug Benefit (ODB) plan in the six months before the index date (i.e., 4494 for Randomization Cohort and 10196 for the Trial cohort). CABG/PCI = Coronary artery bypass grafting / percutaneous coronary intervention.

4 Ontario residents are eligible for the Ontario Drug Benefit program include people 65 years or older, on social assistance, residing in homes for special care and long-term care homes, people receiving professional home care services, and registrants in the Trillium Drug Program.

5 This is an adapted version of the Charlson comorbidity index explicitly created for use in patients with ESRD. This version has a modified weighting scheme specific to dialysis patients.

3.2 Results from the principal component analysis

We subjected 29 of the 156 baseline characteristics to principal component analysis (Appendix 4b). We retained ten principal components that accounted for 61% of the 29 baseline characteristics variance. Appendix 6 and Appendix 7 show results from the principal component analysis.

3.3 Randomization of hemodialysis centers

Each of the 72 participating centers had an approximately 50% chance of being randomized to either trial arm (see Appendix 8 for the process and hardware specification). We observed that some pairs of centers were allocated to different trial arms at a different probability than we might have expected if we had used simple randomization (Fig. 1A to C). These pairs of centers tended to be large and generally had over 225 patients.

3.4 Balance of baseline characteristics

Table 2 shows the balance for a select set of baseline characteristics by the method of constraining. In the Trial Population, both sets of constrained variables were generally well balanced between the two arms, regardless of the randomization method. The constrained randomizations generally provided a slightly better balance. Appendix 9 shows the percentage of times each of the 156 baseline characteristics (from the Trial Population) were balanced across the 1,000 randomization schemes for the three allocation methods. Table 3 shows a summary of the number of baseline characteristics balanced across randomization schemes. The Trial Population had at least 125 of 156 (80%) balanced...
baseline characteristics in 90% of simple randomization schemes. By comparison, the constrained methods always had slightly more balanced baseline characteristics (at least 85% of the 156 baseline characteristics were balanced in 90% of sampled allocations). Table 3 also shows the median (25th, 75th percentile) number of balanced baseline characteristics across the 1,000 sampled randomization schemes by allocation method. The distributions for the number of balanced baseline characteristics were statistically different for the three allocation methods (p-value < 0.0001).
Table 2
The percentage of times each of the baseline characteristics was balanced across each of the 1000 randomizations schemes in the Trial Population.

| Baseline characteristic | Value | Constrained randomization method |
|-------------------------|-------|----------------------------------|
|                         |       | Unrestricted / Simple | Prognostic baseline characteristics | Principal components |
| Center Size             | Mean ± Standard deviation | 32.9% | 41.8% | 38.7% |
| Composite outcome of CV-related death and major CV-related hospitalization | Rate (per 100 person-year) | 32.5% | 36.2% | 33.5% |
| Age (years)             | Mean ± Standard deviation | 95.3% | 99.8% | 99.2% |
|                         | < 65   | 97.8% | 99.7% | 99.9% |
|                         | 65 to 74 | 100.0% | 100.0% | 100.0% |
|                         | 75 to 84 | 100.0% | 100.0% | 100.0% |
|                         | 85 to 105 | 99.5% | 100.0% | 99.9% |
| Sex                     | Male   | 100.0% | 100.0% | 100.0% |
| Living in a rural area  | Yes    | 63.0% | 84.2% | 65.8% |
| Etiology for End-stage kidney disease | Diabetes | 93.0% | 94.5% | 95.0% |
|                         | Glomerulonephritis/autoimmune diseases | 96.3% | 100.0% | 99.5% |
|                         | Drug-induced nephropathy | 100.0% | 99.9% | 100.0% |
|                         | Polycystic kidney disease | 100.0% | 100.0% | 100.0% |
|                         | Renal vascular disease | 97.5% | 97.6% | 96.7% |
|                         | Other  | 88.3% | 91.9% | 91.6% |
| Race                    | Asian  | 75.0% | 81.3% | 88.1% |
|                         | Black  | 73.4% | 95.9% | 91.9% |
|                         | White  | 45.6% | 64.0% | 90.2% |
|                         | Other  | 56.6% | 65.7% | 77.5% |
|                         | Unknown | 93.2% | 93.7% | 93.6% |
| First dialysis modality | Home hemodialysis | 100.0% | 99.8% | 99.9% |
| Baseline characteristic | Value | Unrestricted / Simple | Prognostic baseline characteristics | Principal components |
|-------------------------|-------|-----------------------|-------------------------------------|----------------------|
|                         |       | 97.8%  | 98.6% | 99.9% |
|                         |       | 97.4%  | 98.7% | 99.8% |
| First vascular access used at dialysis start | AV Graft | 99.9%  | 100.0% | 100.0% |
|                         | Fistula | 98.9%  | 99.1% | 99.4% |
|                         | Catheter | 93.5%  | 96.2% | 99.4% |
|                         | PD Catheter | 98.8%  | 99.0% | 100.0% |
|                         | Unknown | 92.4%  | 93.8% | 94.3% |
| Most recent vascular access before the index date | AV Graft | 98.7%  | 99.8% | 98.9% |
|                         | Fistula | 91.9%  | 94.8% | 97.7% |
|                         | Catheter | 89.9%  | 94.0% | 97.4% |
| Patients 65+ years in ODB in the 6 months prior to index date | Yes | 97.5%  | 99.3% | 99.4% |
| Unique hypertensive drugs six months before the index date | Mean ± Standard deviation | 97.1%  | 99.9% | 99.5% |
| Prescribed hypertensive drugs | Angiotensin-converting enzyme (ACE) inhibitors | 99.4%  | 99.3% | 99.5% |
|                         | Angiotensin II Receptor Blocker | 90.7%  | 96.1% | 96.9% |
|                         | Beta-Blockers | 99.7%  | 100.0% | 99.9% |
|                         | Calcium Channel Blocker | 98.1%  | 100.0% | 99.6% |
|                         | Diuretics | 91.9%  | 97.0% | 95.6% |
| CABG/PCI | Yes | 99.4%  | 99.5% | 100.0% |
| Heart failure | Yes | 96.8%  | 100.0% | 99.8% |
| Diabetes mellitus | Yes | 99.0%  | 100.0% | 100.0% |
| Ischemic Stroke | Yes | 100.0% | 100.0% | 100.0% |
| Baseline characteristic                                      | Value                  | Constrained randomization method |
|-------------------------------------------------------------|------------------------|----------------------------------|
|                                                             |                        | Unrestricted / Simple | Prognostic baseline characteristics | Principal components |
| Lower extremity amputation                                  | Yes                    | 99.9%                           | 100.0%                           | 100.0%               |
| Lung disease (COPD)                                         | Yes                    | 99.0%                           | 99.6%                            | 100.0%               |
| Myocardial infarction                                      | Yes                    | 99.2%                           | 100.0%                           | 100.0%               |
| Major Cancer                                                | Yes                    | 100.0%                          | 100.0%                           | 100.0%               |
| Peripheral vascular disease                                | Yes                    | 90.7%                           | 97.2%                            | 91.4%                |
| Modified Charlson comorbidity Score                         | Mean ± Standard deviation | 96.8%                           | 99.9%                            | 100.0%               |
|                                                             | 2                      | 97.7%                           | 100.0%                           | 100.0%               |
|                                                             | 3                      | 100.0%                          | 100.0%                           | 100.0%               |
|                                                             | 4                      | 100.0%                          | 100.0%                           | 100.0%               |
|                                                             | 5+                     | 98.9%                           | 100.0%                           | 100.0%               |
| Having a kidney transplant prior to index date.             | Yes                    | 100.0%                          | 100.0%                           | 100.0%               |
| Number of hospital admissions in the year before the index date | Mean ± Standard deviation | 93.9%                           | 98.4%                            | 98.4%                |
|                                                             | 0                      | 78.4%                           | 76.4%                            | 81.1%                |
|                                                             | 1 to 3                 | 99.5%                           | 99.6%                            | 99.9%                |
|                                                             | 4 to 6                 | 99.4%                           | 99.6%                            | 99.5%                |
|                                                             | 7 to 9                 | 100.0%                          | 99.9%                            | 100.0%               |
|                                                             | 10+                    | 92.1%                           | 92.0%                            | 94.6%                |
| Long term care facility utilization in the year before the index date | Yes                    | 81.3%                           | 86.6%                            | 86.1%                |
| Time since the first date on dialysis (days)                | Mean ± Standard deviation | 88.1%                           | 94.0%                            | 94.4%                |
Table 3  
Summary of the balanced baseline characteristics for the Trial Population.

| Criteria | Constrained randomization method |
|----------|----------------------------------|
|          | Unconstrained / Simple | Prognostic baseline characteristics | Principal components | P-value** |
| 11 prognostic characteristics ++ | | | | |
| Number of constrained baseline characteristics that were balanced in all 1000 (100%) sampled allocations | 0 of 11 (0%) ¥ | 2 of 11 (18%) | 2 of 11 (18%) | < 0.0001 |
| Number of constrained baseline characteristics that were balanced in at least 950 (95%) sampled allocations | 6 of 11 (55%) | 10 of 11 (91%) | 7 of 11 (64%) | |
| Number of constrained baseline characteristics that were balanced in at least 900 (90%) sampled allocations | 8 of 11 (73%) | 10 of 11 (91%) | 9 of 11 (82%) | |
| Median (25th, 75th percentile) number of baseline characteristics that were balanced across the 1000 selected randomization schemes | 10 (9, 11) * | 11 (10, 11) | 10 (10, 11) | |

29 baseline characteristics used in the principal component analysis**

| Number of constrained baseline characteristics that were balanced in all 1000 (100%) sampled allocations | 8 of 29 (28%) | 12 of 29 (41%) | 12 of 29 (41%) | < 0.0001 |
| Number of constrained baseline characteristics that were balanced in at least 950 (95%) sampled allocations | 19 of 29 (66%) | 23 of 29 (79%) | 25 of 29 (86%) | |
| Number of constrained baseline characteristics that were balanced in at least 900 (90%) sampled allocations | 24 of 29 (83%) | 25 of 29 (86%) | 26 of 29 (90%) | |
| Median (25th, 75th percentile) number of baseline characteristics that were balanced across the 1000 selected randomization schemes | 27 (26, 28) | 28 (27, 28) | 28 (28, 29) | |

All 156 available baseline characteristics

| Number of constrained baseline characteristics that were balanced in all 1000 (100%) sampled allocations | 41 of 156 (26%) | 46 of 156 (28%) | 55 of 156 (35%) | < 0.0001 |
| Number of constrained baseline characteristics that were balanced in at least 950 (95%) sampled allocations | 104 of 156 (67%) | 115 of 156 (74%) | 118 of 156 (76%) | |
### Criteria

| Number of constrained baseline characteristics that were balanced in at least 900 (90%) sampled allocations | Unconstrained / Simple | Prognostic baseline characteristics | Principal components | P-value** |
|---|---|---|---|---|
| 125 of 156 (80%) | 132 of 156 (85%) | 134 of 156 (86%) | |

| Median (25th, 75th percentile) number of baseline characteristics that were balanced across the 1000 selected randomization schemes | Unconstrained / Simple | Prognostic baseline characteristics | Principal components | P-value** |
|---|---|---|---|---|
| 147 (142, 150) | 149 (146, 151) | 150 (147, 151) | |

The Trial Population included patients on hemodialysis as of April 1, 2014, and new patients who started in-center hemodialysis during the three-year follow-up. We conducted simple randomization without any restrictions.

¥ For example, for simple randomization, 2 of the 11 chosen prognostic baseline characteristics were always balanced across 1000 randomly sampled allocation schemes.

*For example, for simple randomization, 500 of 1000 allocation schemes had at least ten balanced baseline characteristics out of the 11 prognostic baseline characteristics. As such, there is a 50% probability that a randomly selected allocation will have at least 10 of the 11 prognostic baseline characteristics balanced and a 75% probability that at least 9 of the 11 prognostic baseline characteristics will be balanced.

++ Prognostic baseline characteristics: Constraining on a set of baseline characteristics that thought to be important a priori and included the following patient-level information: age at index date, living in a rural area, Black race, Modified Charlson comorbidity index, number of hospital visits in the previous 12 months, number of unique drugs the patient was prescribed in the six months before the index date, as well as history in the last five years of diagnosis for peripheral vascular disease, congestive heart failure, coronary artery disease, myocardial infarction, number of nephrology consults in the previous 12 months before the index date.

** Results are shown for the 29 baseline characteristics included in the principal component analysis. We did not include any cluster-level baseline characteristics in the constraining process.

### 3.5 Cardiovascular-related death or major cardiovascular-related hospitalization

We followed patients for an average of 1.7 years, and there were 2,260 events over the three-year follow-up. The event rate of the primary outcome was 11 per 100 person-years. Table 4 shows the results from the unadjusted and adjusted analyses for simple and covariate-constrained randomization methods. Across the 1,000 simple randomization schemes for the Trial Population, the mean unadjusted hazard ratio (2.5th, 97.5th percentile) was 1.01 (0.87, 1.16), and 5.9% of allocation schemes produced a confidence interval for the hazard ratio that did not contain the null value of 1.00. Compared to simple randomizations, constrained randomizations had similar unadjusted hazard ratios, with slightly narrower
95% confidence intervals. The Type I error tended to be somewhat lower than the nominal level for some constrained methods than the unconstrained approach.
Table 4
Mean hazard ratio (2.5th, 97.5th percentile) for the composite outcome during a 3-year follow-up of patients on in-center hemodialysis.

| Baseline characteristics adjusted in the analysis | Mean HR (2.5th, 97.5th percentile) | Width of CI | Type 1 error* |
|-------------------------------------------------|-----------------------------------|-------------|---------------|
| **Unadjusted analyses**                          |                                   |             |               |
| Simple (i.e., unconstrained) randomization       | 1.01 (0.87, 1.16)                 | 0.280       | 5.9%          |
| Constrained on a minimal set of baseline         | 1.00 (0.89, 1.12)                 | 0.233       | 3.2%          |
| characteristics Y                               |                                   |             |               |
| Constrained on a minimal set of baseline         | 1.00 (0.88, 1.13)                 | 0.250       | 4.4%          |
| characteristics Y and historical rate of         |                                   |             |               |
| the primary outcome                              |                                   |             |               |
| Constrained on a minimal set of baseline         | 1.00 (0.88, 1.14)                 | 0.260       | 5.2%          |
| characteristics Y and cluster size at time of    |                                   |             |               |
| randomization                                    |                                   |             |               |
| Constrained on a minimal set of baseline         | 1.00 (0.88, 1.13)                 | 0.247       | 4.5%          |
| characteristics Y, historical rate of the         |                                   |             |               |
| primary outcome, and cluster size at time of     |                                   |             |               |
| randomization                                    |                                   |             |               |
| Constrained on 10 principal components           | 1.01 (0.89, 1.12)                 | 0.234       | 3.3%          |
| Constrained on 10 principal components and       | 1.00 (0.88, 1.14)                 | 0.261       | 5.2%          |
| historical rate of primary outcome               |                                   |             |               |
| Constrained on 10 principal components and       | 1.00 (0.87, 1.14)                 | 0.264       | 4.1%          |
| cluster size at time of randomization            |                                   |             |               |
| Constrained on ten principal components, the      | 1.00 (0.89, 1.13)                 | 0.239       | 3.1%          |
| historical rate of the primary outcome, and      |                                   |             |               |
| cluster size at time of randomization            |                                   |             |               |
| **Adjusted for constrained baseline              |                                   |             |               |
| characteristics**                               |                                   |             |               |
| Constrained on a minimal set of baseline         | 1.00 (0.89, 1.12)                 | 0.232       | 8.6%          |
| characteristics Y                               |                                   |             |               |
| Constrained on a minimal set of baseline         | 1.00 (0.89, 1.12)                 | 0.223       | 8.3%          |
| characteristics Y and historical rate of the      |                                   |             |               |
| primary outcome                                  |                                   |             |               |
| Constrained on a minimal set of baseline         | 1.00 (0.89, 1.11)                 | 0.221       | 9.8%          |
| characteristics Y and cluster size at time of     |                                   |             |               |
| randomization                                    |                                   |             |               |
| Constrained on a minimal set of baseline         | 1.00 (0.90, 1.11)                 | 0.216       | 9.6%          |
| characteristics Y, historical rate of the         |                                   |             |               |
| primary outcome, and cluster size at time of     |                                   |             |               |
| randomization                                    |                                   |             |               |
| Constrained on 10 principal components           | 1.00 (0.90, 1.11)                 | 0.203       | 5.2%          |
Baseline characteristics adjusted in the analysis

| Baseline characteristics adjusted in the analysis                                                                 | Mean HR (2.5th, 97.5th percentile) | Width of CI | Type 1 error* |
|-----------------------------------------------------------------------------------------------------------------------|------------------------------------|-------------|---------------|
| Constrained on 10 principal components and historic rate of primary outcome                                          | 1.00 (0.90, 1.11)                 | 0.201       | 6.0%          |
| Constrained on 10 principal components and cluster size at time of randomization                                   | 1.00 (0.90, 1.11)                 | 0.203       | 6.3%          |
| Constrained on ten principal components, the historical rate of the primary outcome, and cluster size at time of randomization | 1.00 (0.91, 1.11)                 | 0.201       | 6.4%          |

All randomization methods had 1000 randomization schemes. The cohort included patients on dialysis as of April 1, 2014, and any patient who started in-center hemodialysis at one of the 72 participating centers during the three-year follow-up.

HR = Hazard ratio; Width of CI = width of confidence interval (i.e., upper minus lower confidence limit).

* The confidence interval's width may not be equal to the difference between the lower and upper confidence limits because of rounding.

Y Included patient-level information: age, living in a rural area, Black race, Modified Charlson comorbidity index, number of hospital visits in the previous 12 months, number of unique drugs the patient was prescribed in the six months before the index date, as well as history in the last five years of diagnosis for peripheral vascular disease, congestive heart failure, coronary artery disease, myocardial infarction, number of nephrology consults in the previous 12 months before the index date.

* Type 1 error in the various constrained scenarios. Note: The nominal Type 1 error is 5%. The observed Type 1 error was within an "acceptable range" if it fell within the 95% confidence interval of the nominal value (i.e., between 3.6% and 6.4%).

** Adjusted analyses included baseline characteristics used in the constraining process.

An acceptable Type 1 error was observed for this method (i.e., between 3.6% and 6.4%).

Adjusted analyses for the constrained methods produced narrower confidence intervals than the unadjusted analyses. However, the Type I error was within the acceptable range only when models adjusted for the ten principal components; the Type I error was outside the expected range for all other adjusted analyses. We also explored the results when adjusting for aggregate-level baseline characteristics as used in the randomization, which aligned with the results when we adjusted for individual-level variables (results not shown).

4. Discussion

This empirical study presented an example of using historical data to conduct covariate-constrained randomization that balances baseline characteristics for a parallel, two-group, cluster randomized trial. Compared to simple randomization, we showed that constraining the random allocation using a historical
cohort (i.e., a Population for Randomization) provides a better balance on baseline characteristics. However, we randomized a moderate number of clusters, and the magnitude of benefit was modest. Our results also suggested that model-based adjustment for the constrained variables produced treatment effects with the nominal Type I error that is narrower than those produced with simple randomization. However, researchers should constrain prognostic variables and adjust for the constrained variables at the analysis stage; otherwise, the Type I error might deviate from the nominal level described in previous reports.\textsuperscript{1,9,11,12,17,18}

In a review of 300 CRTs published between 2000 and 2008, Wright et al.\textsuperscript{49} found significant discrepancies between the restricted randomization used at the design stage and covariate adjustments at the analysis stage. Wright et al.\textsuperscript{49} identified 174 CRTs that used design-based restricted randomization. However, only 30 (17.2\%) of these studies reported an adjusted analysis for all the constrained variables.

From an analysis perspective, the analysis should account for the design that uses covariate-constrained randomization.\textsuperscript{1,9,11,12} Otherwise, the Type I error may deviate from the nominal level because clusters with highly correlated outcomes get separated into different treatment arms (as observed in Fig. 1B and C).\textsuperscript{9} Splitting correlated clusters into different treatment arms tends to (i) lower the Type I error below the nominal level (in the unadjusted analyses), and (ii) decrease power slightly, although we might still expect substantial gains in power due to the assurance of balance on prognostic baseline characteristics.\textsuperscript{9,50} Several analytical techniques can test for treatment effects and take into account the study design. These methods include mixed-effects models, bias-corrected generalized estimating equations, and randomization-based permutation tests.

In our motivating example, we used an analysis for the time-to-first event. In contrast, previous studies have focused their investigations primarily on continuous or binary outcomes.\textsuperscript{1,9,11,12} Our results add to this literature showing a generalized estimating equation-based approach can yield results that maintain the nominal Type I error after adjusting for the covariate-constrained design.

This study has some limitations. First, the Trial Population included a large percentage of patients (37\%) included in the Population for Randomization. Thus, our results may not apply to other designs, for example, CRTs where the Population for Randomization and the Trial Population are the same or settings where cluster- and patient-level profiles change rapidly over time. Second, some historical data may lag by more than one year; thus, these results may not be applicable for Populations at Randomization less than or more than a year old.

Third, our example cohort randomized a moderately large number of clusters; a previous review reported that 75\% of published CRTs randomized fewer than 52 clusters. Covariate-constrained randomization may provide a better baseline balance compared to simple randomization when there are fewer clusters. Finally, our secondary objective does not constitute a formal test of the Type I error. Computer simulations with more control over the generated data would be better suited. As such, the reader should interpret these results cautiously.
5. Conclusions And Guidance For Future Trials

Although covariate-constrained randomization approaches used in this setting had modest improvement for balance, there may be substantial improvements in statistical power.\textsuperscript{12} We propose the following recommendations (Box 1) for CRTs based on the empirical comparisons presented in this paper and other published literature. It is worth noting these recommendations are based on a single setting, and while we anticipate similar findings in different contexts, a more formal statistical comparison would be beneficial.

1. Identify prognostic variables \textit{a priori} using background literature, historical data, or previous trials.
   Previous work for individual-level randomized controlled trials showed increases in statistical power when analyses prespecified covariates strongly associated with the outcome. The adjusted covariates had a more considerable impact on statistical power when the prevalence was moderate to high (between 10% and 50%).\textsuperscript{19,51−53}

2. Researchers should consider generating all (or at least 1,000) simple randomizations to identify baseline characteristics that are always or almost always balanced (e.g.,>95% of the time) between treatment arms. There would be no need to include these baseline characteristics in the constraining process; however, researchers can have these variables in the model-based adjustment to improve the estimates' precision. Importantly, all prognostic variables should be specified \textit{a priori}.\textsuperscript{53}

3. Carefully consider the number of baseline characteristics used during the constraining process. Evidence from our study (and previous simulation studies) showed that over-constraining could result in clusters with highly correlated outcomes having a lower probability of being included in the same trial arm. Thus, over-constraining can lead to a Type I error below the nominal level and slightly decrease power.\textsuperscript{9,50}

4. Researchers can use a dimensionality-reduction method (e.g., principal component analysis) to reduce many dimensions of the prognostic variables to several criterion variables used in the constrained randomization process.\textsuperscript{54} As above, all analyses should account for the dimensionality-reduction criterion at the analytic stage.

5. While the constraining process utilizes aggregate patient-level and cluster-level data, investigators should consider missingness when constraining the randomization on these variables. When appropriate, variables with missing data should be imputed before aggregating the variable at the cluster level.\textsuperscript{55}

6. Researchers should consider constraining the randomization space to the 10% best allocations. Furthermore, researchers should enumerate all possible randomization schemes when fewer than 20 clusters or at least 100,000 randomization schemes.\textsuperscript{1}
Box 1: Guidance for conducting covariate-constrained randomization.

1. Identify prognostic baseline characteristics *a priori* using background literature, historical data, or previous trials.
2. Generate all (or at least 1,000) simple randomizations to identify baseline characteristics that are always balanced between treatment arms (e.g., ≥95% of the time).
3. Carefully consider the number of variables added to the constraining process or consider using a dimensionality-reduction method for many variables (e.g., principal component analysis).
4. Consider the amount of missingness of constrained baseline characteristics prior to randomization.
5. Enumerate all possible allocation schemes when there are fewer than 20 clusters or at least 100,000 allocations otherwise.

Declarations

**Ethics approval and consent to participate**

We had the authorization to use data in this project under section 45 of Ontario’s Personal Health Information Protection Act, which does not require a research ethics board review. The dataset from this study is held securely in coded form at ICES. ICES is an independent, non-profit research institute whose legal status under Ontario’s health information privacy law allows it to collect and analyze healthcare and demographic data, without consent, for health system evaluation and improvement.

**Consent for publication**

Not applicable.

**Availability of data and materials**
While data sharing agreements prohibit ICES from making our study dataset publicly available, access may be granted to those who meet prespecified criteria for confidential access, available at [www.ices.on.ca/das](http://www.ices.on.ca/das). The full dataset creation plan and underlying analytic code can be requested from the authors on the understanding that the computer programs may rely upon coding templates or macros that are unique to ICES and are therefore either inaccessible or may require modification.

**Competing interests**

Authors have nothing to disclose.

**Funding**

Ahmed Al-Jaishi was supported by the Allied Health Doctoral Fellowship from the Kidney Foundation of Canada, CIHR Doctoral Award, and McMaster University Michael DeGroote Scholarship. Stephanie Dixon’s research is supported through a SPOR Innovative Clinical Trial Multi-Year Grant (Grant number: MYG-151209) from CIHR. Amit Garg was supported by the Dr. Adam Linton Chair in Kidney Health Analytics and a Clinician Investigator Award from the CIHR.

We received funding for this study from partnering organizations, including Lawson Health Research Institute, Ontario Renal Network, Dialysis Clinic Inc., Heart and Stroke Foundation of Canada, and Canadian Institutes of Health Research (CIHR) Innovative Clinical Trials Initiative (Grant number: MYG-151209). Funding was also provided by the Ontario Strategy for Patient-Oriented Research SUPPORT Unit, which is supported by the Canadian Institutes of Health Research and the Province of Ontario. Funding bodies had no role in the study's design, analysis, interpretation of data, or writing the manuscript.

**Authors' contributions**

AAA and AXG conceived and led the study design. SND, EM, PJD, and LT contributed to the study design. AA was responsible for data management and analysis. AA drafted the manuscript. All authors contributed to manuscript revision and approved the final manuscript. AA is the guarantor.

**Acknowledgments**

This study was supported by ICES, which is funded by an annual grant from the Ontario Ministry of Health and Long-Term Care (MOHLTC). Parts of this material are based on data and information compiled and provided by CIHI, Cancer Care Ontario (CCO), MOHLTC, and Ontario Service Reports. We thank IMS Brogan Inc. for use of their Drug Information Database. The analyses, conclusions, opinions, and statements expressed herein are solely those of the authors and do not reflect those of the data sources; no endorsement is intended or should be inferred.

**References**

1. Hayes RJ, Moulton LH. *Cluster Randomised Trials*. CRC Press; 2009.
2. Eldridge S, Kerry SM. *A Practical Guide to Cluster Randomised Trials in Health Services Research*. John Wiley & Sons; 2012.

3. Perry M, Faes M, Reelick MF, Olde Rikkert MGM, Borm GF. Studywise minimization: A treatment allocation method that improves balance among treatment groups and makes allocation unpredictable. *J Clin Epidemiol*. 2010;63(10):1118–22. doi:10.1016/j.jclinepi.2009.11.014.

4. Chu R, Walter SD, Guyatt G, et al. Assessment and Implication of Prognostic Imbalance in Randomized Controlled Trials with a Binary Outcome – A Simulation Study. Gong Y, ed. *PLoS One*. 2012;7(5):e36677. doi:10.1371/journal.pone.0036677.

5. Ivers NM, Taljaard M, Dixon S, et al. Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: review of random sample of 300 trials, 2000-8. BMJ. 2011;343(sep26 1):d5886–6. doi:10.1136/bmj.d5886.

6. Raab GM, Butcher I. Balance in cluster randomized trials. *Stat Med*. 2001;20(3):351–365. doi:10.1002/1097-0258(20010215)20:3<351::AID-SIM797>3.0.CO;2-C.

7. Carter BR, Hood K, Fisher R, et al. Balance algorithm for cluster randomized trials. *BMC Med Res Methodol*. 2008;8(1):65. doi:10.1186/1471-2288-8-65.

8. Ivers NM, Halperin IJ, Barnsley J, et al. Allocation techniques for balance at baseline in cluster randomized trials: a methodological review. *Trials*. 2012;13(1):120. doi:10.1186/1745-6215-13-120.

9. Moulton LH. Covariate-based constrained randomization of group-randomized trials. *Clin Trials*. 2004;1(3):297–305. doi:10.1191/1740774504cn024oa.

10. Xiao L, Lavori PW, Wilson SR, Ma J. Comparison of dynamic block randomization and minimization in randomized trials: A simulation study. *Clin Trials*. 2011;8(1):59–69. doi:10.1177/1740774510391683.

11. Li F, Lokhnygina Y, Murray DM, Heagerty PJ, DeLong ER. An evaluation of constrained randomization for the design and analysis of group-randomized trials. *Stat Med*. 2016;35(10):1565–79. doi:10.1002/sim.6813.

12. Li F, Turner EL, Heagerty PJ, Murray DM, Vollmer WM, DeLong ER. An evaluation of constrained randomization for the design and analysis of group-randomized trials with binary outcomes. *Stat Med*. 2017;36(24):3791–806. doi:10.1002/sim.7410.

13. Yu H, Li F, Gallis JA, Turner EL. cvcrand: A Package for Covariate-constrained Randomization and the Clustered Permutation Test for Cluster Randomized Trials. *R J*. 2019;11(2):1–14.

14. Dickinson LM, Beaty B, Fox C, et al. Pragmatic Cluster Randomized Trials Using Covariate Constrained Randomization: A Method for Practice-based Research Networks (PBRNs). *J Am Board Fam Med*. 2015;28(5):663–72. doi:10.3122/jabfm.2015.05.150001.

15. Al-Jaishi AA, McIntyre CW, Sontrop JM, et al. Major Outcomes With Personalized Dialysate TEMPerature (MyTEMP): Rationale and Design of a Pragmatic, Registry-Based, Cluster Randomized Controlled Trial. *Can J Kidney Heal Dis*. 2020;7:1–18. doi:10.1177/2054358119887988.

16. Dempsey AF, Pyrznawoski J, Lockhart S, et al. Effect of a health care professional communication training intervention on adolescent human papillomavirus vaccination a cluster randomized clinical
17. Ford I, Norrie J, Ahmadi S. Model inconsistency, illustrated by the cox proportional hazards model. Stat Med. 1995;14(8):735–46. doi:10.1002/sim.4780140804.

18. Ford I, Norrie J. The role of covariates in estimating treatment effects and risk in long-term clinical trials. Stat Med. 2002;21(19):2899–908. doi:10.1002/sim.1294.

19. Kahan BC, Jairath V, Doré CJ, Morris TP. The risks and rewards of covariate adjustment in randomized trials: An assessment of 12 outcomes from 8 studies. Trials. 2014;15(1):139. doi:10.1186/1745-6215-15-139.

20. Outcomes of a Higher Versus Lower Hemodialysate Magnesium Concentration. Accessed January 20. 2020. https://clinicaltrials.gov/ct2/show/NCT04079582.

21. HiLo | A Pragmatic Clinical Trial. Accessed January 20. 2020. https://hilostudy.org/.

22. Donner A, Klar N. Design and Analysis of Cluster Randomization Trials in Health Research. (Gooster L, Ueberberg A, eds.). Arnold; 2000.

23. Statistics Canada. Population estimates, quarterly. Published August 12, 2020. Accessed August 12, 2020. https://www150.statcan.gc.ca/t1/tbl1/en/tv.action?pid=1710000901.

24. Webster G, Wu J, Williams B, Ivis F, de Sa E, Hall N. Canadian Organ Replacement Register Annual Report: Treatment of End-Stage Organ Failure in Canada 2003–2012.; 2014.

25. Moist LM, Trpeski L, Na Y, Lok CE. Increased Hemodialysis Catheter Use in Canada and Associated Mortality Risk: Data from the Canadian Organ Replacement Registry 2001–2004. Clin J Am Soc Nephrol. 2008;3(6):1726–32. doi:10.2215/CJN.01240308.

26. Ellwood AD, Jassal SV, Suri RS, Clark WF, Na Y, Moist LM. Early Dialysis Initiation and Rates and Timing of Withdrawal From Dialysis in Canada. Clin J Am Soc Nephrol. 2012;8:1–6. doi:10.2215/CJN.01000112.

27. Saczynski JS, Andrade SE, Harrold LR, et al. A systematic review of validated methods for identifying heart failure using administrative data. Pharmacoepidemiol Drug Saf. 2012;21(Suppl 1):129–40. doi:10.1002/pds.2313.

28. Pladevall M, Goff DC, Nichaman MZ, et al. An assessment of the validity of ICD Code 410 to identify hospital admissions for myocardial infarction: The Corpus Christi Heart Project. Int J Epidemiol. 1996;25(5):948–52.

29. Tamariz L, Harkins T, Nair V. A systematic review of validated methods for identifying ventricular arrhythmias using administrative and claims data. Pharmacoepidemiol Drug Saf. 2012;21(Suppl 1):148–53. doi:10.1002/pds.2340.

30. Moist LM, Richards HA, Miskulin D, et al. A Validation study of the Canadian Organ Replacement Register. Clin J Am Soc Nephrol. 2011;6(4):813–8. doi:10.2215/CJN.06680810.

31. Oliver MJ, Quinn RR, Garg AX, Kim SJ, Wald R, Paterson JM. Likelihood of Starting Dialysis after Incident Fistula Creation. Clin J Am Soc Nephrol. 2012;7(3):466–71. doi:10.2215/CJN.08920811.
32. Perl J, Wald R, McFarlane P, et al. Hemodialysis Vascular Access Modifies the Association between Dialysis Modality and Survival. J Am Soc Nephrol. 2011;22(6):1113–21. doi:10.1681/ASN.2010111155.

33. Quinn RR, Laupacis A, Austin PPC, et al. Using administrative datasets to study outcomes in dialysis patients: a validation study. Med Care. 2010;48(8):745–50. doi:10.1097/MLR.0b013e3181e419fd.

34. Al-Jaishi AA, Moist LM, Oliver MJ, et al. Validity of administrative database code algorithms to identify vascular access placement, surgical revisions, and secondary patency. J Vasc Access Published online March. 2018;12:112972981876200. doi:10.1177/1129729818762008.

35. Schultz SE, Rothwell DM, Chen; Z, Tu ; K, Chen Z, Tu K. Identifying cases of congestive heart failure from administrative data: a validation study using primary care patient records. Chronic Dis Inj Can. 2013;33(3):160–6.

36. Hennessy S, Leonard CE, Freeman CP, et al. Validation of diagnostic codes for outpatient-originating sudden cardiac death and ventricular arrhythmia in Medicaid and Medicare claims data. Pharmacoepidemiol Drug Saf. 2010;19(6):555–62. doi:10.1002/pds.1869.

37. Hussain MA, Mamdani M, Saposnik G, et al. Validation of Carotid Artery Revascularization Coding in Ontario Health Administrative Databases. Clin Investig Med Médecine Clin Exp. 2016;39(2):E73-8.

38. Longenecker JC, Coresh J, Klag MJ, et al. Validation of comorbid conditions on the end-stage renal disease medical evidence report: the CHOICE study. Choices for Healthy Outcomes in Caring for ESRD. J Am Soc Nephrol. 2000;11(3):520–9.

39. ICES. Privacy at ICES. Accessed November 25, 2019. https://www.ices.on.ca/Data-and-Privacy/Privacy-at-ICES.

40. Burton A, Altman DG, Royston P, Holder RL. The design of simulation studies in medical statistics. Stat Med. 2006;25(24):4279–92.

41. Greene EJ. A SAS Macro for Covariate-Constrained Randomization of General Cluster-Randomized and Unstratified Designs. J Stat Softw. 2017;77. doi:10.18637/jss.v077.c01.

42. Austin PC. Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research. Commun Stat Simul Comput. 2009;38(6):1228–34. doi:10.1080/03610910902859574.

43. Mamdani M, Sykora K, Li P, et al. Reader’s guide to critical appraisal of cohort studies: 2. Assessing potential for confounding. BMJ. 2005;330:960–2.

44. Yang D, Dalton JE. A unified approach to measuring the effect size between two groups using SAS®. Pap 335–2012 Present 2012 SAS Glob Forum. Published online 2012:1–6.

45. Kruskal WH, Wallis A. Use of Ranks in One-Criterion Variance Analysis. J Am Stat Assoc. 1952;47(260):583–621.

46. Lin DY. Cox regression analysis of multivariate failure time data: The marginal approach. Stat Med. 1994;13(21):2233–47. doi:10.1002/sim.4780132105.

47. Wicklin R. Simulating Data with SAS®. SAS Institute Inc.; 2013.
Tables

Table 1: Select baseline characteristics. The Population for Randomization included patients on hemodialysis as of April 1, 2013. The Trial Population included an open cohort of patients receiving in-center hemodialysis on April 1, 2014, or began receiving in-center hemodialysis during the trial period between April 1, 2014, and March 31, 2017.
| Variable Description                                                                 | VALUE                                      | Population for Randomization | Trial Population |
|--------------------------------------------------------------------------------------|--------------------------------------------|-----------------------------|-----------------|
| Number of centers (n Patients)                                                        | 72 (n=5,812)                               | 72 (n=11,832)               |
| Mean ± Standard deviation                                                             | 81 (69)                                   | 164 (137)                   |
| Historic rate per 100 person-years (cluster standard deviation) (years)               | 10 (3.7)                                  | 11 (3.3)                    |
| Mean ± Standard deviation                                                             | 67 (15)                                   | 66 (15)                     |
| Male                                                                                 | 3373 (58%)                                 | 7069 (60%)                  |
| Diagnoses for End-stage kidney disease                                               |                                            |                             |
| Diabetes                                                                             | 2194 (38%)                                 | 4472 (38%)                  |
| Glomerulonephritis/autoimmune diseases                                                | 882 (15%)                                 | 1575 (13%)                  |
| Drug-induced nephropathy                                                              | 83 (1%)                                   | 159 (1%)                    |
| Polycystic kidney disease                                                             | 229 (4%)                                   | 426 (4%)                    |
| Renal vascular disease                                                                | 1115 (19%)                                 | 1852 (16%)                  |
| Other                                                                                | 738 (13%)                                  | 2221 (19%)                  |
| Unknown                                                                              | 571 (10%)                                  | 1127 (10%)                  |
| Ethnicity                                                                             |                                            |                             |
| Asian                                                                                | 475 (8%)                                   | 854 (7%)                    |
| Black                                                                                | 562 (10%)                                  | 1022 (9%)                   |
| White                                                                                | 3698 (64%)                                 | 7598 (64%)                  |
| Other                                                                                | 1038 (18%)                                 | 2173 (18%)                  |
| Unknown                                                                              | 39 (1%)                                    | 185 (2%)                    |
| Dialysis modality                                                                    |                                            |                             |
| Home hemodialysis                                                                     | 33 (1%)                                    | 84 (1%)                     |
| In-center hemodialysis                                                                | 5215 (90%)                                 | 10529 (89%)                 |
| Peritoneal dialysis                                                                   | 564 (10%)                                  | 1219 (10%)                  |
| Arteriovenous graft                                                                   | 103 (2%)                                   | 198 (2%)                    |
| Arteriovenous fistula                                                                 | 1044 (18%)                                 | 1999 (17%)                  |
| Central venous catheter                                                               | 3927 (68%)                                 | 8157 (69%)                  |
| Peritoneal catheter                                                                   | 456 (8%)                                   | 1092 (9%)                   |
| Vascular access used before the index date                                            |                                            |                             |
| Arteriovenous graft                                                                   | 223 (4%)                                   | 376 (3%)                    |
| Arteriovenous fistula                                                                 | 2159 (37%)                                 | 3461 (29%)                  |
| Central venous catheter                                                               | 3376 (58%)                                 | 7544 (64%)                  |
| Vascular access in the six months                                                     |                                            |                             |
| Yes                                                                                   | 4494 (86%)                                 | 10196 (86%)                 |
| Missing data since the index date | Mean ± Standard deviation |
|-----------------------------------|---------------------------|
| Number of unique hypertensive prescriptions in six months before the index date | 2 (2) 2 (2) |
| Prescribed hypertensive agents | Mean ± Standard deviation |
| Angiotensin-converting enzyme (ACE) inhibitors | 1157 (26%) 2199 (22%) |
| Angiotensin II Receptor Blocker | 1258 (28%) 2413 (24%) |
| Beta-Blockers | 2649 (59%) 5604 (55%) |
| Calcium Channel Blocker | 2392 (53%) 5453 (53%) |
| Diuretics | 1612 (36%) 4242 (42%) |
| CABG/PCI | Yes 1234 (21%) 2612 (22%) |
| Coronary Artery Disease (angina) | Yes 3541 (61%) 6861 (58%) |
| Heart failure | Yes 2862 (49%) 6177 (52%) |
| Diabetes mellitus | Yes 3402 (59%) 7244 (61%) |
| Hypertension | Yes 1528 (26%) 3161 (27%) |
| Ischemic Stroke | Yes 261 (4%) 551 (5%) |
| Lower extremity amputation | Yes 298 (5%) 554 (5%) |
| Chronic obstructive pulmonary disease (COPD) | Yes 2276 (39%) 4666 (39%) |
| Cardiac infarction | Yes 1454 (25%) 2909 (25%) |
| Cancer | Yes 841 (14%) 1803 (15%) |
| Peripheral vascular disease | Yes 1699 (29%) 3055 (26%) |
| Modified Charlson morbidity score | Mean ± Standard deviation |
| 1 | 4 (2) 4 (2) |
| 2 | 1931 (33%) 3631 (31%) |
| 3 | 519 (9%) 1028 (9%) |
| 4 | 1382 (24%) 2883 (24%) |
| 5+ | 1980 (34%) 4290 (36%) |
| Receiving a kidney transplant since the index date | Yes 71 (1%) 148 (1%) |
| Number of days spent in short-term care in the year before the index date | Mean ± Standard deviation |
| 1 (1) | 37 (164) 25 (133) |
| Number of days spent in the hospital in the year before the index date | Mean ± Standard deviation |
| 1 (1) | 11 (28) 17 (31) |
| Number of hospital admissions | Mean ± Standard deviation |
| 1 (1) | 1 (1) |
1 Population for Randomization included patients that were on hemodialysis as of April 1, 2013 index date. The Trial Population included patients on hemodialysis as of April 1, 2014, and any patient who started in-center hemodialysis at one of the 72 participating centers during the three-year trial period. Follow-up ended March 31, 2017. The index date was the first date patients entered the respective cohort.

2 The composite outcome of cardiovascular-related death or hospitalization for myocardial infarction, ischemic stroke, congestive heart failure.

3 Percentages presented only for patients eligible to receive the Ontario Drug Benefit (ODB) plan in the six months before the index date (i.e., 4494 for Randomization Cohort and 10196 for the Trial cohort). CABG/PCI = Coronary artery bypass grafting / percutaneous coronary intervention.

4 Ontario residents are eligible for the Ontario Drug Benefit program include people 65 years or older, on social assistance, residing in homes for special care and long-term care homes, people receiving professional home care services, and registrants in the Trillium Drug Program.

5 This is an adapted version of the Charlson comorbidity index explicitly created for use in patients with ESRD. This version has a modified weighting scheme specific to dialysis patients.
The percentage of times each of the baseline characteristics was balanced across each of the 1000 randomizations schemes in the Trial Population.

| Baseline Characteristic | Value | Unrestricted / Simple | Prognostic baseline characteristics | Principal components |
|-------------------------|-------|-----------------------|-------------------------------------|-----------------------|
| Inter Size              | Mean ± Standard deviation | 32.9%                  | 41.8%                               | 38.7%                 |
| Rate (per 100 person-year) | Unrestricted / Simple | 32.5%                  | 36.2%                               | 33.5%                 |
| Age (years) | Mean ± Standard deviation | 95.3%                  | 99.8%                               | 99.2%                 |
|                       | < 65 | 97.8%                  | 99.7%                               | 99.9%                 |
|                       | 65 to 74 | 100.0%                | 100.0%                              | 100.0%                |
|                       | 75 to 84 | 100.0%                | 100.0%                              | 100.0%                |
|                       | 85 to 105 | 99.5%                 | 100.0%                              | 99.9%                 |
| Gender | Male | 100.0%                  | 100.0%                              | 100.0%                |
| Living in a Rural Area | Yes | 63.0%                  | 84.2%                               | 65.8%                 |
| Cause for End-stage Kidney Disease | Diabetes | 93.0%                  | 94.5%                               | 95.0%                 |
|                       | Glomerulonephritis/autoimmune diseases | 96.3%                  | 100.0%                              | 99.5%                 |
|                       | Drug-induced nephropathy | 100.0%                  | 99.9%                               | 100.0%                |
|                       | Polycystic kidney disease | 100.0%                  | 100.0%                              | 100.0%                |
|                       | Renal vascular disease | 97.5%                  | 97.6%                               | 96.7%                 |
|                       | Other | 88.3%                  | 91.9%                               | 91.6%                 |
| Race | Asian | 75.0%                  | 81.3%                               | 88.1%                 |
|                       | Black | 73.4%                  | 95.9%                               | 91.9%                 |
|                       | White | 45.6%                  | 64.0%                               | 90.2%                 |
|                       | Other | 56.6%                  | 65.7%                               | 77.5%                 |
|                       | Unknown | 93.2%                  | 93.7%                               | 93.6%                 |
| First Dialysis Modality | Home hemodialysis | 100.0%                  | 99.8%                               | 99.9%                 |
|                       | In-center hemodialysis | 97.8%                  | 98.6%                               | 99.9%                 |
|                       | Peritoneal dialysis | 97.4%                  | 98.7%                               | 99.8%                 |
| First Vascular Access Used at Start of Dialysis | AV Graft | 99.9%                  | 100.0%                              | 100.0%                |
|                       | Fistula | 98.9%                  | 99.1%                               | 99.4%                 |
|                       | Catheter | 93.5%                  | 96.2%                               | 99.4%                 |
|                       | PD Catheter | 98.8%                  | 99.0%                               | 100.0%                |
|                                      | Unknown  | 92.4%   | 93.8%   | 94.3%   |
|--------------------------------------|----------|---------|---------|---------|
| AV Graft                             | 98.7%    | 99.8%   | 98.9%   |
| Fistula                              | 91.9%    | 94.8%   | 97.7%   |
| Catheter                             | 89.9%    | 94.0%   | 97.4%   |
| **most recent vascular access before index date** |          |         |         |         |
| Fistula                              | 97.5%    | 99.3%   | 99.4%   |
| **Mean ± Standard deviation**        |          | 97.1%   | 99.9%   | 99.5%   |
| **Mean ± Standard deviation**        |          | 97.1%   | 99.9%   | 99.5%   |
| Angiotensin-converting enzyme (ACE) inhibitors | 99.4%    | 99.3%   | 99.5%   |
| Angiotensin II Receptor Blocker      | 90.7%    | 96.1%   | 96.9%   |
| Beta-Blockers                        | 99.7%    | 100.0%  | 99.9%   |
| Calcium Channel Blocker              | 98.1%    | 100.0%  | 99.6%   |
| Diuretics                            | 91.9%    | 97.0%   | 95.6%   |
| **ABG/PCI**                          | Yes      | 99.4%   | 99.5%   | 100.0%  |
| Heart failure                        | Yes      | 96.8%   | 100.0%  | 99.8%   |
| Diabetes                             | Yes      | 99.0%   | 100.0%  | 100.0%  |
| Renal disease                        | Yes      | 100.0%  | 100.0%  | 100.0%  |
| Vascular occlusion                   |          |         |         |         |
| Stroke                               | Yes      | 99.9%   | 100.0%  | 100.0%  |
| Heart failure                        | Yes      | 99.0%   | 99.6%   | 100.0%  |
| Peripheral vascular disease          | Yes      | 100.0%  | 100.0%  | 100.0%  |
| Major Cancer                         | Yes      | 100.0%  | 100.0%  | 100.0%  |
| Modified Charlson morbidity index    | Mean ± Standard deviation | 96.8%   | 99.9%   | 100.0%  |
| Modified Charlson morbidity index    | 2        | 97.7%   | 100.0%  | 100.0%  |
| Modified Charlson morbidity index    | 3        | 100.0%  | 100.0%  | 100.0%  |
| Modified Charlson morbidity index    | 4        | 100.0%  | 100.0%  | 100.0%  |
| Modified Charlson morbidity index    | 5+       | 98.9%   | 100.0%  | 100.0%  |
| Vascular aneurysm                     | Yes      | 100.0%  | 100.0%  | 100.0%  |
| Number of Hospital Missions in the Year Before the Index Date | Mean ± Standard Deviation | Trial 1 | Trial 2 | Trial 3 |
|-------------------------------------------------------------|---------------------------|---------|---------|---------|
| 0                                                           | 78.4%                     | 76.4%   | 81.1%   |
| 1 to 3                                                      | 99.5%                     | 99.6%   | 99.9%   |
| 4 to 6                                                      | 99.4%                     | 99.6%   | 99.5%   |
| 7 to 9                                                      | 100.0%                    | 99.9%   | 100.0%  |
| 10+                                                         | 92.1%                     | 92.0%   | 94.6%   |

| Long-term Hospitalization in the Year Before the Index Date | Mean ± Standard Deviation | Trial 1 | Trial 2 | Trial 3 |
|-----------------------------------------------------------|---------------------------|---------|---------|---------|
| Yes                                                       | 81.3%                     | 86.6%   | 86.1%   |

**Table 3:** Summary of the balanced baseline characteristics for the Trial Population.
| Criteria                                                                 | Unconstrained / Simple | Prognostic baseline characteristics | Principal components | P-value** |
|------------------------------------------------------------------------|------------------------|-------------------------------------|-----------------------|-----------|
| **11 prognostic characteristics ++**                                  |                        |                                     |                       | <0.0001   |
| Number of constrained baseline characteristics that were balanced in all 1000 (100%) sampled allocations | 0 of 11 (0%) 🅰️       | 2 of 11 (18%)                       | 2 of 11 (18%)         |           |
| Number of constrained baseline characteristics that were balanced in at least 950 (95%) sampled allocations | 6 of 11 (55%)          | 10 of 11 (91%)                      | 7 of 11 (64%)         |           |
| Number of constrained baseline characteristics that were balanced in at least 900 (90%) sampled allocations | 8 of 11 (73%)          | 10 of 11 (91%)                      | 9 of 11 (82%)         |           |
| Median (25th, 75th percentile) number of baseline characteristics that were balanced across the 1000 selected randomization schemes | 10 (9, 11) *           | 11 (10, 11)                         | 10 (10, 11)           |           |
| **29 baseline characteristics used in the principal component analysis** |                        |                                     |                       | <0.0001   |
| Number of constrained baseline characteristics that were balanced in all 1000 (100%) sampled allocations | 8 of 29 (28%)          | 12 of 29 (41%)                      | 12 of 29 (41%)        |           |
| Number of constrained baseline characteristics that were balanced in at least 950 (95%) sampled allocations | 19 of 29 (66%)         | 23 of 29 (79%)                      | 25 of 29 (86%)        |           |
| Number of constrained baseline characteristics that were balanced in at least 900 (90%) sampled allocations | 24 of 29 (83%)         | 25 of 29 (86%)                      | 26 of 29 (90%)        |           |
| Median (25th, 75th percentile) number of baseline characteristics that were balanced across the 1000 selected randomization schemes | 27 (26, 28)            | 28 (27, 28)                         | 28 (28, 29)           |           |
| **All 156 available baseline characteristics**                        |                        |                                     |                       | <0.0001   |
| Number of constrained baseline characteristics that were balanced in all 1000 (100%) sampled allocations | 41 of 156 (26%)        | 46 of 156 (28%)                     | 55 of 156 (35%)       |           |
| Number of constrained baseline characteristics that were balanced in at least 950 (95%) sampled allocations | 104 of 156 (67%)       | 115 of 156 (74%)                    | 118 of 156 (76%)      |           |
| Number of constrained baseline characteristics that were balanced in all 1000 (100%) sampled allocations | 125 of 156 (80%)       | 132 of 156 (85%)                    | 134 of 156 (86%)      |           |
were balanced in at least 900 (90%) sampled allocations

| Median (25th, 75th percentile) number of baseline characteristics that were balanced across the 1000 selected randomization schemes | 147 (142, 150) | 149 (146, 151) | 150 (147, 151) |

The Trial Population included patients on hemodialysis as of April 1, 2014, and new patients who started in-center hemodialysis during the three-year follow-up. We conducted simple randomization without any restrictions.

¥ For example, for simple randomization, 2 of the 11 chosen prognostic baseline characteristics were always balanced across 1000 randomly sampled allocation schemes.

* For example, for simple randomization, 500 of 1000 allocation schemes had at least ten balanced baseline characteristics out of the 11 prognostic baseline characteristics. As such, there is a 50% probability that a randomly selected allocation will have at least 10 of the 11 prognostic baseline characteristics balanced and a 75% probability that at least 9 of the 11 prognostic baseline characteristics will be balanced.

** We used the Kruskal-Wallis H Test to determine whether the three randomization techniques' median number of balanced characteristics was the same.

++ Prognostic baseline characteristics: Constraining on a set of baseline characteristics that thought to be important a priori and included the following patient-level information: age at index date, living in a rural area, Black race, Modified Charlson comorbidity index, number of hospital visits in the previous 12 months, number of unique drugs the patient was prescribed in the six months before the index date, as well as history in the last five years of diagnosis for peripheral vascular disease, congestive heart failure, coronary artery disease, myocardial infarction, number of nephrology consults in the previous 12 months before the index date.

** Results are shown for the 29 baseline characteristics included in the principal component analysis. We did not include any cluster-level baseline characteristics in the constraining process.

Table 4: Mean hazard ratio (2.5th, 97.5th percentile) for the composite outcome during a 3-year follow-up of patients on in-center hemodialysis.
| Baseline characteristics adjusted in the analysis | Mean HR (2.5th, 97.5th percentile) | Width of CI | Type 1 error* |
|-------------------------------------------------|------------------------------------|-------------|---------------|
| **Unadjusted analyses**                         |                                    |             |               |
| Simple (i.e., unconstrained) randomization      | 1.01 (0.87, 1.16)                 | 0.280       | 5.9%          |
| Constrained on a minimal set of baseline...    | 1.00 (0.89, 1.12)                 | 0.233       | 3.2%          |
| and historical rate of the primary outcome     |                                    |             |               |
| Constrained on a minimal set of baseline...    | 1.00 (0.88, 1.13)                 | 0.250       | 4.4%          |
| cluster size at time of randomization          |                                    |             |               |
| Constrained on a minimal set of baseline...    | 1.00 (0.88, 1.14)                 | 0.260       | 5.2%          |
| historical rate of the primary outcome, and... |                                    |             |               |
| cluster size at time of randomization          |                                    |             |               |
| Constrained on 10 principal components         | 1.01 (0.89, 1.12)                 | 0.234       | 3.3%          |
| Constrained on 10 principal components and...  | 1.00 (0.88, 1.14)                 | 0.261       | 5.2%          |
| historical rate of primary outcome            |                                    |             |               |
| Constrained on 10 principal components and...  | 1.00 (0.87, 1.14)                 | 0.264       | 4.1%          |
| cluster size at time of randomization          |                                    |             |               |
| Constrained on ten principal components, the...| 1.00 (0.89, 1.13)                 | 0.239       | 3.1%          |
| historical rate of the primary outcome, and... |                                    |             |               |
| cluster size at time of randomization          |                                    |             |               |
| **Adjusted for constrained baseline characteristics** |                                    |             |               |
| Constrained on a minimal set of baseline...    | 1.00 (0.89, 1.12)                 | 0.232       | 8.6%          |
| and historical rate of the primary outcome     |                                    |             |               |
| Constrained on a minimal set of baseline...    | 1.00 (0.89, 1.12)                 | 0.223       | 8.3%          |
| cluster size at time of randomization          |                                    |             |               |
| Constrained on a minimal set of baseline...    | 1.00 (0.89, 1.11)                 | 0.221       | 9.8%          |
| historical rate of the primary outcome, and... |                                    |             |               |
| cluster size at time of randomization          |                                    |             |               |
| Constrained on 10 principal components         | 1.00 (0.90, 1.11)                 | 0.203       | 5.2%          |
| Constrained on 10 principal components and...  | 1.00 (0.90, 1.11)                 | 0.201       | 6.0%          |
| historical rate of primary outcome            |                                    |             |               |
| Constrained on 10 principal components and...  | 1.00                             | 0.203       | 6.3%          |
Constrained on ten principal components, the historical rate of the primary outcome, and cluster size at time of randomization

|                             | (0.90, 1.11) | 1.00 | (0.91, 1.11) | 0.201 | 6.4% |
|-----------------------------|--------------|------|--------------|-------|------|

All randomization methods had 1000 randomization schemes. The cohort included patients on dialysis as of April 1, 2014, and any patient who started in-center hemodialysis at one of the 72 participating centers during the three-year follow-up.

HR= Hazard ratio; Width of CI = width of confidence interval (i.e., upper minus lower confidence limit).

* The confidence interval's width may not be equal to the difference between the lower and upper confidence limits because of rounding.

** Included patient-level information: age, living in a rural area, Black race, Modified Charlson comorbidity index, number of hospital visits in the previous 12 months, number of unique drugs the patient was prescribed in the six months before the index date, as well as history in the last five years of diagnosis for peripheral vascular disease, congestive heart failure, coronary artery disease, myocardial infarction, number of nephrology consults in the previous 12 months before the index date.

* Type 1 error in the various constrained scenarios. Note: The nominal Type 1 error is 5%. The observed Type 1 error was within an "acceptable range" if it fell within the 95% confidence interval of the nominal value (i.e., between 3.6% and 6.4%).

** Adjusted analyses included baseline characteristics used in the constraining process.

* An acceptable Type 1 error was observed for this method (i.e., between 3.6% and 6.4%).