Contents lists available at ScienceDirect





journal homepage: http://www.elsevier.com/locate/conctc



Covariate-constrained randomization for cluster randomized trials in the long-term care setting: Application to the TRAIN-AD trial



Michele L. Shaffer^{a, b, *}, Erika M.C. D'Agata^c, Daniel Habtemariam^d, Susan L. Mitchell^{d, e}

^a Department of Public Health Sciences, Penn State College of Medicine, Hershey, PA, USA

^b Frank Statistical Consulting LLC, Vashon, WA, USA

^c Division of Infectious Diseases, Warren Alpert Medical School, Brown University, Providence, RI, USA

^d Hebrew SeniorLife Institute Hinda and Arthus Institute for Aging Research, Boston, MA, USA

^e Department of Medicine, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston, MA, USA

ARTICLE INFO

ABSTRACT

Keywords: Covariate-constrained randomization Minimization Cluster randomized trial Study design Little has been reported on strategies to ensure key covariate balance in cluster randomized trials in the nursing home setting. Facilities vary widely on key characteristics, small numbers may be randomized, and staggered enrollment is often necessary. A covariate-constrained algorithm was used to randomize facilities in the Trial to Reduce Antimicrobial use In Nursing home residents with Alzheimer's Disease and other Dementias (TRAIN-AD), an ongoing trial in Boston-area facilities (14 facilities/arm). Publicly available 2015 LTCfocus. org data were leveraged to inform the distribution of key facility-level covariates. The algorithm was applied in waves (2-8 facilities/wave) June 2017-March 2019. To examine the algorithm's general performance, simulations calculated an imbalance score (minimum 0) for similar trial designs. The algorithm provided good balance for profit status (Arm 1, 7 facilities; Arm 2, 6 facilities). Arm 2 was allocated more nursing homes with the number of severely cognitive impaired residents above the median (Arm 1, 7 facilities; Arm 2, 10 facilities), resulting in an imbalance in total number of residents enrolled (Arm 1, 196 residents; Arm 2, 228 residents). Facilities with number of black residents above the median were balanced (7 facilities/arm), while the numbers of black residents enrolled differed slightly between arms (Arm 1, 26 residents (13%); Arm 2, 22 residents (10%)). Simulations showed the median imbalance for TRAIN-AD's original randomization scheme (score = 3), was similar to the observed imbalance (score = 4). Covariate-constrained randomization flexibly accommodates logistical complexities of cluster trials in the nursing home setting, where LTCfocus.org is a valuable source of baseline data.

Trial registration number and trial register: ClinicalTrials.gov Identifier: NCT03244917.

1. Introduction

In cluster randomized clinical trials (RCTs), balancing covariates between arms at both the individual and cluster level is important. Despite the growing number of cluster RCTs in nursing homes (NHs), relatively little has been reported on strategies to ensure balance of key covariates.

Ivers et al. [1] reviewed allocation methods for cluster RCTs. Stratification and matching are common approaches to restricted randomization. However, stratification can balance only a few covariates. Matching is capable of achieving balance for multiple covariates, but it is dependent on optimal matches being identified through an intense screening process, and may be inefficient if the intra-cluster correlation is low. Further, if one cluster within a match is lost to followup, the pair of clusters is lost for analysis. Thus, as in small clinical trials randomized at the individual level [2], methods beyond stratification and matching are needed to ensure balance.

Minimization is another widely used method of intervention allocation and is effective in achieving balance for multiple covariates with relatively small numbers of clusters [3]. For individually randomized trials, minimization randomly assigns the first participant to an arm. Subsequent participants are assigned by selecting the alloca-

https://doi.org/10.1016/j.conctc.2020.100558

Received 25 November 2019; Received in revised form 3 March 2020; Accepted 14 March 2020

Available online 17 March 2020

^{*} Corresponding author. 90 Hope Drive, Suite 2200, Hershey, PA, 17033, USA.

E-mail addresses: mls177@psu.edu, mls177@psu.edu (M.L. Shaffer), edagata@lifespan.org (E.M.C. D'Agata), DanielHabtemariam@hsl.harvard.edu (D. Habtemariam), smitchell@hsl.harvard.edu (S.L. Mitchell).

^{2451-8654/© 2020} The Author(s). Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licensex/by-nc-nd/4.0/).

tion which minimizes imbalance, while accounting for the attributes of participants previously allocated. Minimization can similarly apply to cluster randomization. While minimization is most often conducted using categorized covariates, methods are available for the inclusion of continuous covariates [4]. A limitation of minimization is the predictability of assignment beyond the first cluster, which could lead to selection bias. To overcome this limitation, an element of randomness can be added [5].

When baseline information is available, a more complex version of randomization can be used. Covariate-constrained randomization enumerates all possible allocations of participating clusters. Next, this list of allocations is limited to those that minimize imbalance across baseline covariates. Trial allocation is chosen randomly from this limited list. Thus, the final allocation minimizes imbalance while retaining an element of randomness. Modified [6] and blocked [7] versions of covariate-constrained randomization have been proposed in other settings.

The application of covariate-constrained randomization to cluster RCTs in the NH setting has not been described but is particularly compelling for several reasons. First, NHs tend to be varied in characteristics that can affect outcomes such as size, profit status, and demographic characteristics [8,9]. Second, NH interventions are often multi-modal and complex, and implementation tends to necessitate staggered recruitment and randomization of facilities in waves over time. Lastly, for cluster RCTs with relatively fewer clusters, covariate-constrained randomization may achieve better balance than minimization, while being less vulnerable to selection bias [1].

We describe a covariate-constrained randomization algorithm applied to the ongoing National Institutes of Health funded cluster RCT entitled Trial to Reduce Antimicrobial use In Nursing home residents with Alzheimer's Disease and other Dementias (TRAIN-AD) [10]. Data from LTCfocus.org [11] are used to provide baseline information for facility-level covariates prior to allocation. We further examine the algorithm's average and worst-case performance in simulation studies for designs similar to TRAIN-AD to inform its future use.

2. Methods

The conduct of the TRAIN-AD study was approved by the Institutional Review Board (IRB) at Hebrew SeniorLife.

2.1. LTCfocus.org data

We used the LTCfocus.org data [11] as a source of baseline information to inform our covariate constrained randomization of NHs. LTCfocus.org is part of the Shaping Long-Term Care in America Project conducted at the Brown University Center for Gerontology and Healthcare Research [11]. The publicly available database annually compiles nationwide data describing facility features in all federally licensed US NHs. Sources include Certification and Survey Provider Enhanced Reporting (CASPER) [12], which are facility-level data such as aggregated resident characteristics and nursing home organizational characteristics; the Area Health Resource Files [13], which contain county-level data about health professionals and facilities; state level Medicaid policy data [14] such as payment rates, reimbursement methodology, and bed hold policies; and the Minimum Data Set (MDS) [15], which includes diagnoses, treatments, medications, and activities of daily living for residents aggregated to the facility, county, and state levels.

2.2. TRAIN-AD design overview

TRAIN-AD is a cluster RCT funded by the National Institutes of Health. A detailed description of the protocol is published elsewhere [10]. Briefly, the trial is evaluating a multicomponent antimicrobial

stewardship program to improve antimicrobial prescribing in the management of suspected infections with advanced dementia by merging best practices in palliative care and infectious diseases. The intervention is implemented at the NH level, primarily by providing training to clinical providers (nurses, physicians, nurse practitioners and physician assistants) caring for advanced dementia residents. Training is delivered through multiple modalities including: in-person seminars, an on-line course, management algorithms, and structured feedback of prescribing practices. The control NHs employ usual care. In all NHs, residents with advanced dementia are identified and enrolled at the time the NH first entered the trial and every two months thereafter up to 12 months. The medical charts of these residents are abstracted every two months to determine the occurrence and management of suspected urinary tract and respiratory tract infections. Each resident is followed for up to 12 months or until death. The primary trial outcome, the number of antimicrobials courses for these suspected infections/person-day alive over 12 months, is hypothesized to be lower in the intervention arm. As this is deemed a minimal risk study and all data are already collected for clinical purposes, individual informed consent is waived by the IRB, and thus virtually all eligible residents are enrolled.

2.3. Facility recruitment and randomization

Eligible NHs had to have greater than 60 beds and be located within 60 miles of Boston as ascertained from the LTCfocus.org 2015 data [11]. Study information was mailed to the senior administrators of facilities meeting these criteria, who were telephoned one week later by a research team member to solicit their facility's participation, including their agreement to randomize the facility to either the intervention or control arm. As designed, the trial was to include 24 Boston-area NHs (N = 12/arm) staggered in four waves of six facilities.

Allocation in waves was chosen as it is very challenging to have all NHs agree in advance of the study start date to participate, and impossible from the perspective of the implementation team to onboard all intervention NHs at one time. Each intervention home required a three-month start-up period with substantial work on the part of the research team. If all NHs were recruited in advance, many would need to wait months to a year to enter the study. Considering the high turnover of senior administrators that characterizes the NH industry, a dynamic recruitment strategy for facilities was critical.

Once the trial was initiated, the number of NHs was increased to a total of 28 NHs (N = 14/arm) to meet the target sample size estimates for residents (N = 410 residents, N = 205/arm). In addition, due to four facility drop-outs that were later replaced, the number of NHs in each wave varied, and the total number of NHs randomized was 32. Thus, ultimately, NHs were recruited, randomized, and initiated into the trial in six staggered waves approximately 3 months apart beginning in June 2017 and ending March 2019. The following number of NHs recruited and randomized per wave: Wave 1, 6 NHs (3 control, 3 intervention); Wave 2, 6 NHs (3 control, 3 intervention); Wave 5, 2 NHs (1 control, 1 intervention); and Wave 6, 2 NHs (1 control, 1 intervention).

Prior to any recruitment efforts, key characteristics of Boston-area facilities (N = 95) meeting the aforementioned eligibility criteria were ascertained using LTCfocus.org 2015 data [11]. Characteristics on which to base the covariate-constrained randomization algorithm were selected based on prior research reporting key NH variables associated with advanced dementia care [8,9], including: profit status, the number of residents with advanced cognitive impairment in the facility, and the number of black residents in the facility.

The primary analysis planned for TRAIN-AD is a resident-level analysis. Because the number of residents enrolled can vary substantially from NH to NH using our opt-out approach, we sought to keep the overall sample size balanced between groups using the approximate number of residents with severe cognitive deficit as a proxy for eligible residents. The distribution of the number of black residents across all eligible NHs is really semi-continuous with many zeroes and skewed right. In addition, since the number of black residents is only a proxy for the number of black residents with advanced dementia we decided to use a coarser measure splitting at the median. A very similar categorization results from splitting the percentages of black residents at the median.

The number of residents with advanced cognitive impairment reflects an attempt to balance sample size in the two arms. To estimate the number of residents with advanced cognitive impairment, the number of beds in the facility was multiplied by the percent of residents with a Cognitive Function Scale (CFS) of 3 (range, 0–3, with 3 indicating severe cognitive impairment) [16]. A similar approach was used to estimate the number of black residents. Facilities were then categorized as having higher or lower representation of severely cognitive impaired and black residents based on median values of all eligible facilities. It was crucial to balance on race as it is one of the strongest factors influencing end-of-life care in advanced dementia, and NHs in the United States tend to be highly segregated [17].

A waved version of covariate-constrained randomization was initially applied for the original study design, which planned for four waves of six facilities (three to control, three to intervention) per wave, but then was adapted to accommodate facility dropouts and the increase in number of enrolled NHs. For each wave, we used a generalization of the method of minimization for allocation [1,7]. We selected this method to preserve desirable elements of covariateconstrained randomization, while providing flexibility to replace or add additional facilities if needed. For the first wave of six facilities, all possible ways of assigning three facilities to each arm were considered, and one of the multiple assignments that minimized imbalances across the three covariates was selected randomly. The imbalance score for the first wave was the sum of the absolute differences between the number of NHs in each arm that were for-profit status, above the median for number of residents with severe cognitive impairment, and above the median for number of black residents. For subsequent waves, all potential assignments of the new facilities were considered, and the assignments from previous waves were included in calculating the cumulative imbalance score. When multiple assignments yielded equivalently best imbalance scores, an assignment was chosen randomly. If facilities dropped out prior to the initiation of data collection, they were replaced in a subsequent wave, and the imbalance scores were updated to account for the characteristics of facilities that dropped out. The statistician for the trial (MLS) was masked to the identity of facilities as well as the actual facility assignments. Facilities were identified using codes in the list of eligible NHs with their attributes from the LTCfocus.org data [11]. Allocation assignments were returned in a partially masked form as "A" or "B" to the project director. The project director arbitrarily assigned A and B to intervention and control.

2.4. Statistical analysis

We tabulated the key characteristics of each NH recruited for TRAIN-AD including the profit status, number of residents enrolled, and number of black residents enrolled as well as other characteristics of the facilities and residents important for analysis and interpretation of the study. Other facility characteristics of interest included bed size, registered nurse hours per resident day, and five-star rating score; resident characteristics included age, gender, and Bedford Alzheimer Nursing Severity-Subscale score (BANS-S) [18]. The balance of these other characteristics served as a randomness check for our constrained algorithm. We calculated frequencies and percentages or medians and ranges to compare the distributions of the actual NHs recruited with the information available from the 2015 LTCfocus.org data [11]. We tested for differences in the study arms using chi-square tests for binary facility-level characteristics and Wilcoxon rank sum tests for quantitative facility-level characteristics. Generalized estimating equations (GEE) were used to compare resident-level characteristics between study arms. We used the Pearson correlation coefficient to quantify correlation and the concordance correlation coefficient [19] to quantify agreement between the number of black residents enrolled and total number of residents enrolled per facility and the corresponding LTCfocus.org facility-level data [11].

We conducted simulations to examine how well our planned TRAIN-AD randomization scheme, as well as alternative schemes, would work for similar studies. Using only one wave of randomization is equivalent to standard covariate-constrained randomization with binary covariates, while the other extreme of randomizing each of the 24 individual NHs separately is similar to minimization. Thus, as alternative randomization schemes, we considered minimization incorporating a random element of 0.8 and scenarios with a smaller and a larger number of waves: covariate-constrained randomization (one wave of 24 NHs), three waves of eight NHs, and six waves of four NHs. We did not include NH dropout as these simulations are intended to reflect planned designs for allocating facilities. For each allocation scheme, we randomly selected 10,000 sets of 24 distinct NHs from the list of 95 facilities eligible prior to the start of study recruitment with each set of NHs representing a hypothetical study. Since each subset of 24 NHs is randomly ordered, we assigned NHs to waves sequentially. For example, for a design with three waves of eight facilities, the first eight nursing homes randomly selected were included in the first wave, the second eight in the next wave, and the final eight in the third wave. For each hypothetical study, we applied the waved minimization algorithm used for TRAIN-AD and calculated the overall imbalance score (minimum 0 with higher scores indicating greater imbalance) in the same manner as for the actual study. Across the 10,000 sets of NHs for each randomization scheme we calculated measures of average performance, the mean and median, and measures of variability and worst-case performance, the standard deviation and maximum, respectively. Due to the large number of possible combinations of homes for covariate-constrained randomization (one wave of 24 NHs) to consider, we randomly selected 10,000 combinations for each of the simulations.

3. Results

Characteristics for the 95 facilities eligible prior to recruitment are shown in Table 1. The majority of NHs (71%) were for profit. There was a large range in both the estimated number of black residents (0– 164) and estimated number of severely cognitive impaired residents (2–106) per facility. The estimated number of black residents was particularly skewed with a median of 3 residents per facility.

Table 1

Characteristics of 95 nursing homes eligible for TRAIN-AD at the start of study recruitment based on 2015 LTCfocus.org data.

Characteristic ^a	Frequency (Percentage) or Median (Minimum- Maximum)
For Profit	67 (71%)
Estimated number of black residents (number of beds	3 (0–164)
multiplied by percent of black residents)	
Estimated number of severely cognitive impaired	22 (2–106)
residents ^a (number of beds multiplied by percent of	
severely cognitive impaired residents)	

^a Residents with a Cognitive Function Scale (CFS) of 3 (range, 0–3, with 3 indicating severe cognitive impairment).

For summarizing imbalance, we report the two arms as Arm 1 and Arm 2, with identification of actual intervention and control masked, as the trial is not yet complete.

Table 2 shows the final balance for the key characteristics we sought to balance as well as other characteristics relevant to analysis and interpretation of the study. Arm 1 was allocated one more forprofit NH than Arm 2 (Arm 1, N = 7 NHs; Arm 2, N = 6 NHs). Arm 2 was allocated three more facilities than Arm 1 (Arm 1, N = 7 NHs; Arm 2, N = 10 NHs) where the number of severely cognitive impaired residents (CFS = 3) was greater than the median based on the LTCfocus.org data [11] and as a result, ultimately, a greater number of residents with advanced dementia was enrolled in Arm 2 (N = 228 residents) compared to Arm 1 (N = 196 residents). Arms 1 and 2 were allocated the same number of facilities (N = 7 NHs/arm) where the number of black residents was greater than the median based on the LTCfocus.org data [11], and the percentage of residents with advanced dementia who were black that were ultimately enrolled into the study was reasonably balanced between arms (13% in Arm 1 (N = 26 residents) vs. 10% in Arm 2 (N = 22 residents)). Thus, the

Table 2

Facility- and resident-level characteristics by intervention arm for TRAIN-AD.

Characteristic	Frequency (%) or Mean [Median]			
	Arm 1	Arm 2	p-Value	
Facility-level	N = 14	N = 14		
For-profit status	7 (50)	6 (43)	0.71	
Number of black residents	1.9 [0]	1.6 [0]	0.81	
Number of residents enrolled	14 [12]	16 [14]	0.41	
Five-star rating score	4.2 [5.0]	3.7 [4.0]	0.39	
Bed Size	139 [137]	173 [130]	0.84	
Registered nurse hours/resident day	1.36 [1.35]	1.25 [1.20]	0.12	
Resident-level	N = 196	N = 228		
Race black	26 (13)	22 (10)	0.67	
Gender female	161 (82)	190 (83)	0.72	
Age	88 [89]	85 [87]	0.09	
BANS-S score ^a	19.6 [19.5]	20.4 [20.5]	0.04	

^a Bedford Alzheimer Nursing Severity-Subscale (BANS-S: range 7–28, higher scores indicate more functional disability).

overall imbalance score for the 28 facilities recruited and retained for TRAIN-AD was 4. Other facility and resident characteristics showed relatively good balance with no large discrepancies between arms observed. While the test of means was statistically significant for BANS-S, the arms differed on average by less than one point.

To further examine sources of imbalance, Figs. 1 and 2 show the estimated numbers from the LTCfocus.org data [11] and the actual enrolled numbers for black residents and all residents with advanced dementia, respectively. While there is positive correlation between the estimated and actual enrolled numbers for both black residents (Pearson r = 0.90) and all residents (Pearson r = 0.54), and the larger correlation for black residents may in part explain why black residents were more balanced between arms than all residents, there are large differences in the absolute values (concordance correlation coefficients 0.17 and 0.21, respectively). These differences are more pronounced for estimated numbers above the medians.

Results of the simulations are reported in Table 3. The mean imbalance score for four waves of six NHs (the original TRAIN-AD randomization scheme) was 3.3 with a median imbalance score of 3 and a maximum imbalance score of 9. The minimum imbalance score for all randomization schemes was 0. As expected, the average imbalance, variability (as measured by the standard deviation), and maximum imbalance increased as the number of waves increased. The maximum imbalance for the waved approaches was lower than randomization using minimization. However, the average performance was similar to minimization for larger numbers of waves.

Fig. 3 shows the distribution of the imbalance score from the 10,000 simulations for the original TRAIN-AD randomization scheme of four waves of six NHs, which is relatively symmetric with a slight skew right. The observed imbalance score of 4 is near the center of the distribution with approximately 42% of the simulations having an imbalance score of 4 or higher.

4. Discussion

We used an adapted covariate-constrained randomization scheme to balance profit status, number of black residents, and number of residents with severe cognitive impairment (as a proxy for eligible resi-



Fig. 1. Estimated number of black residents versus actual number of black residents enrolled for each nursing home. Points are jittered to avoid overlap. The dashed line represents the median estimated number of black residents for all eligible TRAIN-AD facilities at the start of study recruitment. The solid line is the line of agreement (y = x). Intervention arm is partially masked, and reported as Arms 1 and 2.



Fig. 2. Estimated number of residents versus actual number of residents enrolled for each nursing home. Points are jittered to avoid overlap. The dashed line represents the median estimated number of residents for all eligible TRAIN-AD facilities at the start of study recruitment. The solid line is the line of agreement (y = x). Intervention arm is partially masked, and reported as Arms 1 and 2.

Table 3

Results of 10,000 simulations randomizing 24 nursing homes from the list of all eligible facilities for TRAIN-AD at the start of study recruitment using randomization schemes with different numbers of waves.

Operating Characteristic	Randomization Scheme				
	One Wave of 24 Facilities	Three Waves of Eight Facilities	Four Waves of Six Facilities	Six Waves of Four Facilities	Minimization
Median imbalance score	1	3	3	4	3
Mean imbalance score	1.5	2.8	3.3	4.1	3.3
Standard deviation of imbalance score	0.86	1.4	1.5	1.9	1.9
Minimum imbalance score	0	0	0	0	0
Maximum imbalance score	3	8	9	13	18

dents). The ability to use covariate-constrained randomization instead of minimization or simpler forms of restricted randomization was made feasible by the availability of covariate data from LTCfocus.org [11]. The overall balance for the three covariates was modest. We achieved the best possible balance for profit status, given the odd number of for-profit NHs. The numbers of total residents, and to a lesser extent black residents, remained somewhat imbalanced.

The imbalance score was calculated as the sum of the marginal absolute differences in number of facilities between intervention and control arms for the three covariates. While a lower imbalance score is better, there is no specific threshold for the score to determine if sufficient balance is achieved. The score can be viewed as a relative measure, as illustrated when comparing alternative designs in the simulations. This imbalance score is similar to the total imbalance score calculated by the range method for minimization with equal weighting of facility characteristics [20]. A limitation of the score is its prevalence dependence; very low or very high prevalence characteristics are more likely to be balanced by chance. Thus, alternative weights may need to be considered. There are several ways to generalize the score depending upon the application, including weighting covariates in the sum to allow one or more covariates to carry greater importance, and allowing categorical covariates with more than two categories or continuous covariates [1,21].

The adapted covariate-constrained randomization scheme used for TRAIN-AD provided the flexibility to accommodate staggered entry and recruitment. While matching is also possible, the need to identify good matches may have delayed entry of NHs into the study, and the loss of a pair of NHs if one home dropped out would have greatly impacted recruitment. Alternative designs have been suggested to allow staggered entry and recruitment. One such alternative is steppedwedge designs [22-24]. A complication of using a stepped-wedge design in the long-term care setting is that the number of new residents entering a facility over time may be too few to allow for a traditional cross-sectional design to be completed in a reasonable timeframe. A closed-cohort design in which new residents cannot join [25], also may not be feasible, since approximately 40% of residents die when followed over a period of 12 months [26]. Thus, such studies would require an open cohort design [25]. Sample size calculations for both open and closed cohort stepped wedge designs are less developed [27].

NHs show substantial segregation by race, and race greatly influences advanced dementia care [17]. Thus, it is very important to balance race in cluster RCTs in the NH setting. A prior cluster RCT of residents in NHs with advanced dementia used matching to balance profit status alone but this left race imbalanced, requiring adjustment in the primary analysis [28]. A later cluster RCT of residents in NHs with advanced dementia used matching on both profit status and race, but race still showed imbalance [29]. While we were able to balance the number of NHs recruited to each arm that fell above or below the median estimated number of black residents based on the LTCfocus. org data [11], there was a small imbalance in the actual numbers of



Fig. 3. Distribution of imbalance score for 10,000 simulations randomizing 24 nursing homes from the list of all eligible facilities for TRAIN-AD at the start of study recruitment using the original randomization scheme for TRAIN-AD of four waves of six nursing homes.

black resident enrolled in each arm. One explanation is that we balanced higher or lower estimated number of black residents based on the median, rather than using the estimated number of black residents itself. The estimated number of black residents also does not directly measure the number of cognitive impaired black residents that would have been eligible for the trial. Finally, the LTCfocus.org data [11] used were from 2015 and may not accurately reflect current resident populations.

Similarly, we included whether the estimated number of residents with severe cognitive impairment was higher or lower than the median estimated number of residents with severe cognitive impairment among all eligible NHs to balance overall enrollment. Again, the estimated number of residents with severe cognitive impairment does not directly measure the number of severely cognitive impaired residents. Specifically, we note that the CFS score, on which the LTCfocus.org data [11] is based, does not include other factors in our resident eligibility criteria, most notably the diagnosis of dementia. While CFS validates well with the Global Deterioration Scale [16], it is not exactly the same. We noted some large deviations when comparing these estimated numbers to the actual numbers of residents enrolled in each facility (Fig. 2).

We did not update the medians used to separate facilities into lower or higher numbers of black residents and residents with severe cognitive impairment in each wave, as they were not greatly impacted by the addition of the small number of new NHs (N = 8) deemed eligible following the start of study recruitment. The recruitment of additional NHs was ultimately necessary, and the algorithm was easily adapted to include additional waves of NHs to fulfill recruitment. Since we implemented a hybrid form of minimization and covariateconstrained randomization, the process can be applied to waves as small as 2 NHs. However, utilizing only waves with a small number of NHs is not recommended. Particularly with small waves, it is possible only a single best allocation would exist, and thus to avoid assignments that are deterministic a random element that considers other assignments with a small probability should be added. We included a scenario in our simulations for a greater number of waves of fewer NHs to show the general impact of recruiting more waves, where the performance is comparable to minimization.

Covariate-constrained randomization provides a flexible approach to intervention allocation for cluster randomized trials in the longterm care setting. The LTCfocus.org data provide a rich source of baseline data to facilitate covariate-constrained randomization for intervention allocation in future studies in the long-term care setting.

Acknowledgements

This research was supported in part by the National Institutes of Health National Institute on Aging [NIH-NIA R37AG032982, NIH-NIA K24AG033640 (to SLM)] and National Institute of Allergy and Infectious Diseases [NIH-NIAID 1K24AI119158-01A1 (to EMCD)]. The LTCfocus.org data are part of the Shaping Long Term Care in America Project at Brown University funded in part by the National Institute on Aging (1P01AG027296).

References

- N.M. Ivers, I.J. Halperin, J. Barnsley, et al., Allocation techniques for balance at baseline in cluster randomized trials: a methodological review, Trials 13 (2012) 120–6215, https://doi.org/10.1186/1745-6215-13-120.
- [2] D.G. Altman, J.M. Bland, How to randomise, BMJ 319 (7211) (1999) 703–704.
- [3] D.G. Altman, J.M. Bland, Treatment allocation by minimisation, BMJ 330 (7495) (2005) 843 doi: 330/7495/843 [pii].
- [4] S. Senn, V.V. Anisimov, V.V. Fedorov, Comparisons of minimization and Atkinson's algorithm, Stat. Med. 29 (7–8) (2010) 721–730, https://doi.org/10. 1002/sim.3763.
- [5] J. Hofmeijer, P.C. Anema, I. van der Tweel, New algorithm for treatment allocation reduced selection bias and loss of power in small trials, J. Clin. Epidemiol. 61 (2) (2008) 119–124, https://doi.org/10.1016/j.jclinepi.2007.04. 002.
- [6] P.J. Nietert, R.G. Jenkins, L.S. Nemeth, S.M. Ornstein, An application of a modified constrained randomization process to a practice-based cluster randomized trial to improve colorectal cancer screening, Contemp. Clin. Trials 30 (2) (2009) 129–132, https://doi.org/10.1016/j.cct.2008.10.002.
- [7] B.R. Carter, K. Hood, Balance algorithm for cluster randomized trials, BMC Med. Res. Methodol. 8 (2008) 65–2288, https://doi.org/10.1186/1471-2288-8-65.
- [8] S.E. Engel, D.K. Kiely, S.L. Mitchell, Satisfaction with end-of-life care for nursing home residents with advanced dementia, J. Am. Geriatr. Soc. 54 (10) (2006) 1567–1572 doi: JGS900 [pii].
- [9] S.L. Mitchell, J.M. Teno, J. Roy, G. Kabumoto, V. Mor, Clinical and organizational factors associated with feeding tube use among nursing home residents with advanced cognitive impairment, J. Am. Med. Assoc. 290 (1) (2003) 73–80, https://doi.org/10.1001/jama.290.1.73.
- [10] A.J. Loizeau, E.M.C. D'Agata, M.L. Shaffer, et al., The trial to reduce antimicrobial use in nursing home residents with Alzheimer's disease and other

dementias: study protocol for a cluster randomized controlled trial, Trials 20 (1) (2019) 594-019, https://doi.org/10.1186/s13063-019-3675-y.

- [11] Brown University Center for Gerontology and Healthcare Research. Long-term Care, Facts on Care in the US –, LTCfocus.org. Ltcfocus.org.
- [12] Centers for Medicare & Medicaid Services, (2019). https://www.cms.gov/ Medicare/Quality-Initiatives-Patient-Assessment-Instruments/ HomeHealthQualityInits/downloads/HHQICASPER.pdf.
- [13] Health Resources & Services Administration, (2019). https://data.hrsa.gov/ topics/health-workforce/ahrf.
- [14] Centers for Medicare & Medicaid Services, (2019). https://www.cms.gov/ Regulations-and-Guidance/Guidance/Manuals/Paper-Based-Manuals-Items/ CMS021927.html.
- [15] Centers for Medicare & Medicaid Services, (2019). https://www.cms.gov/ research-statistics-data-and-systems/computer-data-and-systems/minimum-dataset-3-0-public-reports/index.html.
- [16] K.S. Thomas, D. Dosa, A. Wysocki, V. Mor, The minimum data set 3.0 cognitive function scale, Med. Care 55 (9) (2017) e68–e72, https://doi.org/10.1097/MLR. 00000000000334.
- [17] D.B. Smith, Z. Feng, M.L. Fennell, J.S. Zinn, V. Mor, Separate and unequal: racial segregation and disparities in quality across U.S. nursing homes, Health Aff (Millwood) 26 (5) (2007) 1448–1458 doi: 26/5/1448 [pii].
- [18] L. Volicer, A.C. Hurley, D.C. Lathi, N.W. Kowall, Measurement of severity in advanced Alzheimer's disease, J. Gerontol. 49 (5) (1994) M223–M226, https:// doi.org/10.1093/geronj/49.5.m223.
- [19] L.I. Lin, A concordance correlation coefficient to evaluate reproducibility, Biometrics 45 (1) (1989) 255–268.
- [20] S.J. Pocock, R. Simon, Sequential treatment assignment with balancing for prognostic factors in the controlled clinical trial, Biometrics 31 (1975) 103–115.
- [21] F. Li, Y. Lokhnygina, D.M. Murray, P.J. Heagerty, E.R. DeLong, An evaluation of

constrained randomization for the design and analysis of group-randomized trials, Stat. Med. 35 (10) (2016) 1565–1579, https://doi.org/10.1002/sim.6813.

- [22] C.A. Brown, R.J. Lilford, The stepped wedge trial design: a systematic review, BMC Med. Res. Methodol. 6 (2006) 54–2288 doi: 1471-2288-6-54 [pii].
- [23] E. de Hoop, I. van der Tweel, R. van der Graaf, et al., The need to balance merits and limitations from different disciplines when considering the stepped wedge cluster randomized trial design, BMC Med. Res. Methodol. 15 (2015) 93–9015, https://doi.org/10.1186/s12874-015-0090-2.
- [24] M.A. Hussey, J.P. Hughes, Design and analysis of stepped wedge cluster randomized trials, Contemp. Clin. Trials 28 (2) (2007) 182–191 doi: S1551-7144 (06)00063-2 [pii].
- [25] K. Hemming, M. Taljaard, J.E. McKenzie, et al., Reporting of stepped wedge cluster randomised trials: extension of the CONSORT 2010 statement with explanation and elaboration, BMJ 363 (2018) k1614, https://doi.org/10.1136/ bmj.k1614.
- [26] S.L. Mitchell, J.M. Teno, D.K. Kiely, et al., The clinical course of advanced dementia, N. Engl. J. Med. 361 (16) (2009) 1529–1538, https://doi.org/10.1056/ NEJMoa0902234.
- [27] K. Hemming, T.P. Haines, P.J. Chilton, A.J. Girling, R.J. Lilford, The stepped wedge cluster randomised trial: rationale, design, analysis, and reporting, BMJ 350 (2015) h391, https://doi.org/10.1136/bmj.h391.
- [28] S.L. Mitchell, M.L. Shaffer, S. Cohen, L.C. Hanson, D. Habtemariam, A.E. Volandes, An advance care planning video decision support tool for nursing home residents with advanced dementia: a cluster randomized clinical trial, JAMA Intern. Med. 178 (7) (2018) 961–969, https://doi.org/10.1001/jamainternmed. 2018.1506.
- [29] L.C. Hanson, S. Zimmerman, M.K. Song, et al., Effect of the goals of care intervention for advanced dementia: a randomized clinical trial, JAMA Intern. Med. 177 (1) (2017) 24–31, https://doi.org/10.1001/jamainternmed.2016.7031.