



Published in final edited form as:

Clin Trials. 2023 December ; 20(6): 613–623. doi:10.1177/17407745231185055.

Assessing the representativeness of cluster randomized trials: Evidence from two large pragmatic trials in United States nursing homes

Nina R Joyce^{1,2}, Sarah E Robertson^{3,4,5}, Ellen McCreedy^{2,3}, Jessica Ogarek^{2,3}, Edward H Davidson⁶, Vincent Mor^{2,3}, Stefan Gravenstein^{2,3,7}, Issa J Dahabreh^{4,5,8}

¹Department of Epidemiology, Brown University School of Public Health, Providence, RI, USA

²Center for Gerontology and Health Care Research, Brown University School of Public Health, Providence RI, USA

³Department of Health Services Policy and Practice, Brown University School of Public Health, Providence, RI, USA

⁴CAUSALab, Harvard T.H. Chan School of Public Health, Boston, MA, USA

⁵Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA

⁶Insight Therapeutics, LLC, Norfolk, VA, USA

⁷Warren Alpert Medical School of Brown University, Providence, RI, USA

⁸Department of Biostatistics, Harvard T.H. Chan School of Public Health, Boston, MA, USA

Abstract

Background/Aims: When the randomized clusters in a cluster randomized trial are selected based on characteristics that influence treatment effectiveness, results from the trial may not be directly applicable to the target population. We used data from two large nursing-home-based pragmatic cluster randomized trials to compare nursing home and resident characteristics in randomized facilities to eligible non-randomized and ineligible facilities.

Methods: We linked data from the High Dose (HD) influenza vaccine trial and the Music & Memory Pragmatic TRIal for Nursing Home Residents with ALzheimer’s Disease (METRICaL) to nursing home assessments and Medicare fee-for-service claims. The target population for the HD trial comprised Medicare certified nursing homes; the target population for the METRICaL trial comprised nursing homes in one of four U.S. based nursing home chains. We used standardized mean differences to compare facility and individual characteristics across the three groups and logistic regression to model the probability of nursing home trial participation.

Author Contributions: NRJ and IJD conceptualized the study; NRJ extracted data, conducted all analyses, created all tables and figures, and wrote the first draft of the manuscript; IJD provided critical methodological input on the original and all subsequent drafts; IJD, SER, EM, JO, VM, SG, EHD critically reviewed the manuscript and provided substantive expertise on interpretation of findings; SG, EHD, and VM provided the data for the manuscript; all authors reviewed and approved the final version for submission.

Trial registration: [ClinicalTrials.gov NCT01815268](https://clinicaltrials.gov/ct2/show/study/NCT01815268), [NCT04850807](https://clinicaltrials.gov/ct2/show/study/NCT04850807), and [NCT03821844](https://clinicaltrials.gov/ct2/show/study/NCT03821844)

Results: In the HD trial, 4,476 (29%) of the 15,502 nursing homes in the target population were eligible for the trial, of which 818 (18%) were randomized. Of the 1,361,122 residents, 91,179 (6.7%) were residents of randomized facilities, 463,703 (34.0%) of eligible non-randomized facilities, and 806,205 (59.3%) of ineligible facilities. In the METRICaL trial, 160 (59%) of the 270 nursing homes in the target population were eligible for the trial, of which 80 (50%) were randomized. Of the 20,262 residents, 973 (34.4%) were residents of randomized facilities, 7,431 (36.7%) of eligible non-randomized facilities, and 5,858 (28.9%) of ineligible facilities. In the HD trial, randomized facilities differed from eligible non-randomized and ineligible facilities by the number of beds (132.5 vs. 145.9 and 91.9, respectively), for-profit status (91.8% vs. 66.8% and 68.8%), belonging to a nursing home chain (85.8% vs. 49.9% and 54.7%), and presence of a special care unit (19.8% vs. 25.9% and 14.4%). In the METRICaL trial randomized facilities differed from eligible non-randomized and ineligible facilities by the number of beds (103.7 vs. 110.5 and 67.0), resource-poor status (4.6% vs. 10.0% and 18.8%), and presence of a special care unit (26.3% vs. 33.8% and 10.9%). In both trials, the characteristics of residents in randomized facilities were similar across the three groups.

Conclusion: In both trials, facility-level characteristics of randomized nursing homes differed considerably from the eligible non-randomized and ineligible facilities, while there was little difference in resident-level characteristics across the three groups. Investigators should assess the characteristics of clusters that participate in cluster randomized trials, not just the individuals within the clusters, when examining the applicability of trial results beyond participating clusters.

Keywords

Nursing home; Medicare; generalizability; influenza; dementia

Background

Over the past three decades there has been a dramatic increase in the conduct of cluster randomized trials.¹ In cluster randomized trials, the unit of randomization is a group of individuals united by a common feature, such as their neighborhood, care provider, or school. Eligible individuals in the cluster are randomly assigned to the same treatment, and outcomes are assessed at the level of the individual.^{1–3} In many cases, clusters can be assumed independent of each other, even though observations within a given cluster may be dependent. For this reason, cluster randomization is often the preferred approach for comparing interventions when the treatment status of one individual may affect the outcome of another (i.e., in the presence of spillover effects/interference), or when the intervention is best implemented at the cluster level.⁴ For example, cluster randomization is frequently used in vaccine trials, where an individual's outcome depends on both their own treatment status and the treatment status of the individuals around them because of herd immunity, and is also common in clinical interventions that target the facilities, sites or providers where individuals affected by the intervention receive services instead of the individual patients.

In a cluster randomized trial, the eligible clusters that are invited and agree to participate in the trial (i.e., the randomized clusters) can be viewed as a (possibly) selected sample from a target population of eligible clusters.⁵ When the randomized clusters are not representative of the target population because they are selected based on characteristics that

influence treatment effectiveness or safety, results from the trial are not directly applicable to the target population. Though several studies have assessed the representativeness of individually randomized trials,^{6–8} much less attention has been paid to the representativeness of cluster randomized trials.⁹ This is true despite surveys of trialists suggesting participating clusters are seldom a random sample from the target population.^{10,11} Instead, these surveys have found that trial participation is often a function of investigator preferences (e.g., selecting clusters to meet logistical needs) and administrative personnel preferences (e.g., in clusters with limited resources participation may be viewed as an added burden with little benefit).

In this paper, we use data from two large pragmatic cluster randomized trials in U.S. nursing homes, one of which was nested within a cohort of all Medicare-accredited U.S. nursing homes and the other of which was nested within the cohort of nursing homes that belong to one of four nursing home chains. We compare the characteristics between randomized, eligible but non-randomized homes, and ineligible homes. Pragmatic trials typically have broad eligibility criteria, and their results are often presumed to generalize to real-world populations. Applying treatment effect estimates from a pragmatic cluster trial to a target population, however, requires the trial sample to be representative of the target population with respect to the distribution of effect modifiers, not just that the trial's eligibility criteria encompass the target population.^{12–15} Despite the pragmatic nature of the design of both trials that we examine here, facility-level characteristics in randomized nursing homes differed considerably from those in the eligible non-randomized and ineligible nursing homes, but there was little variation in resident-level characteristics across the three groups.

Methods

Cluster randomized trials

The High-Dose (HD) trial was a multicenter pragmatic cluster randomized trial conducted in U.S. nursing homes during the 2013–2014 flu season ([NCT01815268](#)) to compare the effectiveness and safety of high-dose versus standard-dose influenza vaccine on the risk of hospitalization, mortality, and functional decline.^{16,17} The trial had a 2×2 factorial design in which the nursing homes were randomized to provide (1) high dose versus standard dose influenza vaccine to eligible residents and (2) standard dose influenza vaccine versus standard of care to nursing home staff. Both the high dose and standard dose vaccines were provided free to nursing home staff.

The Music & Memory Pragmatic TRIal for Nursing Home Residents with Alzheimer's Disease (METRICal) was a pragmatic trial ([NCT04850807](#)) that began in 2019 ([NCT03821844](#)) conducted at 81 nursing homes in the United States owned by one of four corporate chains. The purpose of the trial was to assess the effect of creating and playing tailored music playlists versus standard of care on the frequency of agitated and aggressive behaviors using inexpensive iPod technology to deliver individualized music to people with Alzheimer's disease and related dementias at early signs of agitation. In supplemental table 1 we list eligibility criteria for nursing homes and nursing home residents who were candidates for the intervention in each trial (hereafter referred to as “candidates for the intervention”).

Nursing home and resident data

In both trials, the investigators used routinely updated data from administrative databases (maintained for clinical care and billing purposes) to collect baseline characteristics and outcomes for all trial participants. Specifically, the investigators obtained data on nursing home characteristics from the Long-Term Care Focus database, which compiles information across five different data sources to create an annual “snapshot” of each nursing home.¹⁸ Investigators also obtained data on nursing home residents from the Minimum Data Set (MDS) 3.0, a comprehensive assessment of all nursing home residents conducted upon entry and then every quarter thereafter and at times of significant changes in health status. The MDS contains information on demographics, chronic and acute illnesses, medication use, and procedures done in the nursing home.

Both the HD and METRICaL trials are cluster versions of a nested trial design in which randomized nursing homes are nested within a sample from the target population of eligible nursing homes.¹⁹ The MDS and Long-Term Care Focus database include the same information for all nursing homes and residents regardless of trial participation. Thus, in both trials we used the same sources of data for all nursing homes and residents to assess the representativeness of the randomized facilities to their respective target population.

Target populations

For the HD trial, the target population comprised all U.S. based Medicare certified nursing homes as of October 2013; for the METRICaL trial, the target population comprised U.S. based Medicare certified nursing homes that belonged to one of four chains as of January 2018. In both trials we defined nursing homes that participated in the trial (i.e., were randomized) as the “randomized facilities”, and nursing homes that were eligible for the trial but did not participate as the “eligible non-randomized facilities”.

For both trials, several eligibility criteria were related to practical aspects of planning and conducting the research rather than the health effects of the interventions. Thus, in practice, if the interventions were deemed safe and effective, they would be offered in the ineligible nursing homes as well. Therefore, we also compared nursing home and resident characteristics of randomized nursing homes to ineligible nursing homes. Specifically, in the HD trial we defined the ineligible nursing homes as all Medicare certified nursing homes that did not meet eligibility criteria for the trial; in the METRICaL trial we defined the ineligible nursing homes as all nursing homes that belonged to one of the four chains that did not meet the eligibility criteria.

Facility and resident characteristics

We compared facility-level and resident-level characteristics of the randomized nursing homes to the eligible non-randomized and ineligible nursing homes.⁵ We focused on facility and resident covariates that prior studies and background knowledge suggest are strongly associated with the risk of the outcomes in each trial and thus are likely modifiers of the treatment effect (on the difference or ratio scale).^{20–24} Specifically, at the facility-level, we included operational features of the nursing home (number of beds, occupancy rate, and presence of a special care unit for dementia or other condition), ownership status (owned

by a nursing home chain and not-for-profit tax status; we only used this information for the HD trial because only nursing homes that belonged to one of the four for-profit chains were invited, and agreed, to participate in the METRICaL trial). We also included measures of patient care and need (the number of total direct care hours provided per day and the number of direct care hours provided by registered nurses, certified nursing assistants, and licensed practical nurses per day), as well as the acuity index for the nursing home (a composite measure of the level of need for the residents based on the activities of daily living scale) and the designation as a resource poor facility (a measure reflecting the payment source for residents).

Facility-level covariates were assessed in 2013 for the HD trial and in 2018 for the METRICaL trial (the year before each trial began). In addition, we compared the proportion of residents in the nursing home that were candidates for the intervention. For the HD trial we also compared trial outcomes (all-cause mortality and hospitalization for influenza-like illness) in the year prior to the trial and the proportion of residents who received an influenza vaccine in the year prior to the trial. Outcomes for the METRICaL trial were collected using a specialized survey instrument, the Cohen-Mansfield Agitation Inventory Scale, and thus outcomes were not available outside the trial because they are not collected in the administrative data.

At the resident level, we considered demographics (age, sex, race, marital status), diabetes, Alzheimer's disease, pneumonia, chronic kidney disease, congestive heart failure hypertension, cirrhosis, atrial fibrillation and coronary artery disease, and procedures and treatments done while a resident of the facility (chemotherapy, radiation therapy, dialysis, the use of a ventilator, and the placement of an indwelling urinary catheter). We assessed resident-level comorbidities and procedures on the most recent MDS assessment prior to the start of each trial. We deemed comorbidities to be present if the checkbox for the condition was checked or if the relevant International Classification of Disease (ICD)-9 (HD trial) or ICD-10 (METRICaL trial) code was included as a write-in diagnosis on the MDS data collection instrument (see supplemental table 2 for the ICD-9 and ICD-10 codes used to define each condition). Procedures were identified from MDS indicators only.

Statistical analysis

We estimated standardized mean differences to compare nursing-home-level and resident-level characteristics in the randomized nursing homes to the eligible non-randomized and ineligible nursing homes. We considered an absolute standardized mean difference of greater than 0.1 as an indicator of meaningful differences between the groups of facilities.²⁵

When performing comparisons across facilities, we averaged resident-level covariates at the facility-level (e.g., the proportion of residents with diabetes in the facility) separately for residents who were candidates for treatment and those who were not. For example, we compared the proportion of residents with diabetes who were and were not candidates for treatment separately across the facility groups (randomized, eligible non-randomized, and ineligible).

We used logistic regression at the facility level to model the probability of nursing home participation in each trial among eligible nursing homes, conditional on all facility characteristics and averages of resident-level characteristics included in Table 1 grouped by whether they were candidates for the interventions. We plotted smoothed densities of the estimated probability of trial participation, separately for randomized and eligible non-randomized nursing homes. We assessed model discrimination by graphing the Receiver Operating Characteristic (ROC) curve and estimating the area under the ROC curve (i.e., the c-statistic). Additionally, we fit a multinomial model adjusted for all facility level measures and aggregated resident level measures to assess the association between facility characteristics and groupings of facilities treated (i.e., randomized, eligible non-randomized, and ineligible). All analyses were performed in R version 4.0.2 (R Core Team, 2020, Vienna, Austria).

Results

Randomized, eligible non-randomized, and ineligible nursing homes

Of the 15,502 US based nursing homes in 2013, 4,476 (29%) met eligibility criteria for the HD trial, of which 818 (18% of eligible facilities) were randomized to high or low-dose influenza vaccine, leaving 3,658 eligible non-randomized facilities (24% of all facilities and 82% of all eligible facilities). On October 1, 2013 there were 1,361,122 residents of US based nursing homes that could be assessed for eligibility in the HD trial; 91,179 (6.7%) of whom were residents of randomized facilities, 463,341 (34.0%) of whom were residents of eligible non-randomized facilities, and 806,602 (59.3%) of whom were residents of ineligible facilities.

Of the 270 facilities across the four nursing home chains in the METRICaL trial, 160 (59%) met eligibility criteria for the trial, of which 80 (50% of eligible facilities) were randomized, leaving 80 eligible non-randomized facilities (30% of all homes and 50% of all eligible homes). On January 1st, 2018 there were 20,262 residents across the 270 facilities, of whom 6,973 (34.4%) were residents of randomized facilities, 7,431 (36.7%) were residents of eligible non-randomized facilities, and 5,858 (28.9%) were residents of ineligible facilities.

Facility characteristics

In Tables 1 and 2 we present the nursing home characteristics, and in Figures 1 and 2 we plot the standardized mean differences, comparing randomized facilities to eligible non-randomized and ineligible facilities. In the HD trial, when comparing randomized nursing homes to eligible non-randomized facilities, all facility level measures except for 2012 hospitalization rates had a standardized mean difference greater than 0.1. Specifically, randomized nursing homes had fewer beds (132.5 vs. 145.9), were substantially more likely to be part of a nursing home chain (85.8% vs. 49.9%) and be a for-profit facility (91.8% vs. 66.8%), and were less likely to have any special care unit (19.8% vs. 25.9%), including an Alzheimer's/Dementia unit (17.9% vs. 23.8%). Although there was little difference in hospitalization rates in the year prior to the trial, randomized facilities had a lower rate of influenza vaccination (65.2% vs. 72.1%).

When comparing ineligible facilities (n=11,026) to randomized facilities, all measures except for 2012 hospitalization rate, 2012 mortality rate, and registered nursing hours/person-per-day (PPD) had a standardized mean difference greater than 0.1. Most notably, compared to ineligible facilities, randomized nursing homes had on average more beds (132.5 vs. 91.9), were less likely to be resource poor (5.4% vs. 9.2%), and were more likely to be a for profit facility (91.8% vs. 68.8%).

In the METRICaL trial, when comparing randomized facilities to eligible non-randomized facilities, all facility level measures had a standardized mean difference greater than 0.1 with the exception of registered nurse to staff ratios, direct care hours/PPD, and belonging to nursing home chain B or C. Specifically, randomized nursing homes were substantially less likely to be resource-poor (10.0% vs. 18.8%), had fewer total beds (103.7 vs. 110.5), a greater percentage of the randomized facilities belonged to nursing home chain C (18.8% vs. 6.3%) and a smaller percentage belonged to nursing home chain D (37.5% vs. 50.0%). When comparing ineligible facilities (n=110) to randomized facilities, only direct care hours/PPD, certified nursing assistants/PPD, and occupancy rates had an absolute value of the standardized mean difference less than 0.1. Most notably, compared to ineligible facilities, randomized nursing homes had more beds (110.5 vs. 67.0), were more likely to be part of nursing home chain C (18.8% vs. 1.8%), and were more likely to have any special care unit (33.8% vs. 10.9%).

Resident characteristics

In Tables 1 and 2 we present the resident characteristics, and in Figures 3 and 4 we plot the standardized mean differences, comparing residents of the randomized facilities to residents in the eligible non-randomized and ineligible facilities. In the HD trial, resident characteristics tended to be more similar than facility characteristics between the randomized and eligible non-randomized, with the largest differences observed in the percentage of residents who were candidates for treatment (54.7% vs. 64.8%), and the percentage of residents with a diagnosis of Alzheimer's disease and related dementias (50.0% vs. 58.6%). There were no meaningful differences between the residents of randomized and ineligible facilities.

In the METRICaL trial, there was a small difference in the percent of residents with Alzheimer's disease and related dementias between the residents of randomized and eligible non-randomized facilities (43.1% vs. 48.6%). The largest differences between the residents in randomized and ineligible nursing homes were observed for White race (68.8% vs. 83.0%), age (77.6 vs. 81.7), and a diagnosis of atrial fibrillation (5.2%, vs. 9.1%).

For both the HD and METRICaL trial, when we stratified comparisons by whether residents were candidates for the randomized interventions, our findings were unchanged (Supplemental figures 1–4).

Probability of trial participation

In Supplemental Figure 5 we plot the predicted probability of trial participation among eligible nursing homes for the HD and METRICaL trials.

In the HD trial, the median predicted probability of trial participation was 0.42 (25th percentile = 0.20; 75th percentile = 0.64), and the mean was 0.43 (standard deviation [SD] = 0.26) among the randomized facilities, while the median was 0.07 (25th percentile = 0.03; 75th percentile = 0.18), and the mean was 0.13 (SD = 0.14) among the eligible non-randomized facilities. The c-statistic for this model was 0.84.

In the METRICaL trial, the median predicted probability of trial participation was 0.74 (25th percentile = 0.55; 75th percentile = 0.88) and the mean was 0.70 (SD = 0.23) in the randomized facilities, while in the eligible non-randomized facilities the median was 0.27 (25th percentile = 0.10; 75th percentile = 0.49) and the mean was 0.31 (SD = 0.25). The c-statistic for this model was 0.86.

When we fit the multinomial model to assess the association between facility characteristics and facility grouping (randomized, eligible non-randomized, ineligible), almost all facility level measures were significantly associated with trial eligibility and randomization status (Supplementary Tables 3–4).

Discussion

Pragmatic cluster randomized trials have the potential to provide valuable information on the effectiveness of interventions under “real world” circumstances.²⁶ Yet, when the distribution of effect modifiers in the clusters that participate in the trial differs from the distribution in the target population of clusters where the trial results will be applied, the estimates from the trial may not reflect the effect of the intervention in the target population. In this study, we compared characteristics of nursing homes and their residents in two large pragmatic cluster randomized trials against the characteristics of eligible non-randomized and ineligible nursing homes and their residents. We found little difference in the characteristics of the residents of the facilities, but substantial differences in the characteristics of the participating facilities as compared to the eligible non-randomized and ineligible facilities. That the largest differences were at the cluster, and not the individual level, suggests that if there is an association between nursing homes and the characteristics of their patient population, this does not appear to translate into eligibility for the trial, or participation in the trial among the eligible facilities, for the limited set of variables we were able to measure. Our findings highlight the importance of assessing the representativeness of the clusters that participate in cluster randomized trials, not just the individuals within the clusters, especially when cluster characteristics may modify the effectiveness or safety of treatment.^{27–29}

We included comparisons of the randomized to ineligible facilities for two reasons. First, trial eligibility criteria may reflect logistical and structural aspects of research planning and conduct that are unrelated to the effectiveness and safety of the intervention being studied. For instance, the requirement in the HD trial that nursing homes be within 50 miles of a Centers for Disease Control and Prevention (CDC) reporting city was based on investigator interest in comparing community and nursing home influenza rates, which required the trial facilities to be close to a CDC reporting city. Thus, many of the ineligible facilities were ineligible for logistical, not clinical, reasons that may not be modifiers of the treatment effect. Second, even when there are reasons to suspect that the treatment effect

may differ by the eligibility criteria, as is often the case, it is reasonable to expect that managers of ineligible nursing homes would consider findings from the trials when deciding whether to offer high-dose influenza vaccine to candidate residents. Thus, we included comparisons to ineligible facilities to better reflect the “real world” process of clinical and policy decision-making. Nevertheless, our finding that, in both trials, the randomized facilities were somewhat more similar to the eligible non-randomized facilities than to the ineligible facilities suggests that the differences are not entirely logistical. Investigators may want to consider selecting eligibility criteria that do not modify the treatment effect if they are interested in generalizing inferences beyond the trial.

Future research could attempt to account for differences between the randomized and non-randomized facilities and use formal methods to extend causal inferences from the trials to their respective target populations.^{30–34} When extending inferences to a target population, investigators need rich data on both the clusters participating in the trial and the target population where the treatments will be implemented. For the HD and METRICaL trials, we were able to use data from the Long-Term Care Focus database and MDS 3.0, both of which are comprehensive and commonly used assessments of facilities and their residents. Although data for both trials were collected from the same data source, we conceptualized the target population of each trial differently. Thus, several facility level characteristics were strongly associated with participation in one trial but not associated, or strongly associated with non-participation, in the other (e.g., resource poor status, a special unit, the RN/nurse ratio, number of direct care hours provided by RNs showed different patterns of association between trials). In addition to differences in the measured characteristics associated with trial participation, however, there may be characteristics of the facilities and the residents within the facilities that are strong predictors of the facility’s decision to participate in the trial that are not captured in the datasets. For example, while distance is not an eligibility criterion for the METRICaL trial, in discussions with the trialists, the nursing home directors reported a preference for nursing homes that were close to the chain headquarters to limit the travel time required for staff delivering the intervention (personal communication, Drs. Gravenstein and McCreedy). For the HD trial, trialists indicated preference for participation of multi-facility corporations for operational efficiency of the study, reducing representation of facilities with less organized corporate oversight that could affect infection control practices (personal communication, Drs. Gravenstein and Davidson). Thus, not only did the target populations differ, but so too did the mechanisms for selection into the trial, highlighting the importance of clearly defining the target population and understanding the recruitment process that results in the trial sample and potential for affecting the measured outcomes.

Investigators interested in extending inferences beyond the trial participants should think carefully about which variables that modify the treatment effect (or predict the outcome) are associated with trial participation based on their substantive knowledge in both the design³⁵ and analysis stage of the trial.³⁴ In the design stage investigators may want to build additional data collection and qualitative research into the process to determine the factors driving selection into the trial. For example, investigators can identify “non-standard” measures that may not be routinely collected, such as the preferences of the trialists conducting the trial that may drive participation among the eligible clusters.⁹ In

the analysis stage of the trial, we agree with prior recommendations that the inclusion of a CONSORT diagram specific to cluster randomized trials could greatly improve efforts to extend inferences beyond the trial even when investigators are not involved in the trial design and implementation.^{36,37}

As studies similar to ours improve our understanding of the selection process into cluster randomized trials, their findings may be used to support analyses adjusting for differences using formal statistical methods.^{30–33} Our findings highlight the importance of assessing the characteristics of clusters that participate in cluster randomized trials, not just the individuals within the clusters, when investigators are interested in extending their results beyond participating clusters.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Conflicts of interest:

Grants/Financial Support: This work was supported in part by the Agency for Healthcare Research and Quality (AHRQ) award K12 HS022998-03 (Joyce); National Library of Medicine (NLM) award R01LM013616 (Dahabreh); Patient-Centered Outcomes Research Institute (PCORI) Methods Research Awards ME-1502-27794 and ME-2021C2-22365 (Dahabreh); National Institute on Aging (NIA) award R33AG057451 (Mor); Research grant from Sanofi (Swiftwater, PA, USA) to fund the HD trial (Gravenstein).

Statements in this paper do not necessarily represent the views of the AHRQ, NLM, PCORI, PCORI's Board of Governors, PCORI's Methodology Committee, NIA, or Sanofi.

Mor is a member of the Sanofi Portfolio Advisory Board for which he is paid an honorarium. Dahabreh reports consulting for Moderna on methods for observational analyses. Dahabreh is also the Principal Investigator of a research agreement between Harvard University and Sanofi Pasteur for a pilot study on methods for transportability analyses using individually randomized vaccine trials. Moderna and Sanofi Pasteur did not have any role in the planning or conduct of the research reported in this paper. Gravenstein is a consultant for Sanofi Pasteur, Janssen, Pfizer, GlaxoSmith Kline, Novavax, Vaxart, and receives research funding from Sanofi Pasteur, Pfizer, and Genentech. Davidson has received research funding from Sanofi Pasteur and Genentech.

Mor is a member of the Sanofi Portfolio Advisory Board for which he is paid an honorarium. Dahabreh reports consulting for Moderna on methods for observational analyses. Dahabreh is also the Principal Investigator of a research agreement between Harvard University and Sanofi Pasteur for a pilot study on methods for transportability analyses using individually randomized vaccine trials. Moderna and Sanofi Pasteur did not have any role in the planning or conduct of the research reported in this paper. Gravenstein is a consultant for Sanofi Pasteur, Janssen, Pfizer, GlaxoSmith Kline, Novavax, Vaxart, and receives research funding from Sanofi Pasteur, Pfizer, and Genentech. Davidson has received research funding from Sanofi Pasteur and Genentech.

Abbreviations that appear in the text:

CDC	Centers for Disease Control and Prevention
CI	Confidence Interval
HD	High dose trial
METRICaL	Music & Memory Pragmatic TRIal for Nursing Home Residents with Alzheimer's Disease

LTC Focus	Long Term Care Focus
MDS	Minimum Data Set
PPD	Per-Person-per-Day
ROC	Receiver Operating Characteristic
SD	Standard Deviation

Abbreviations that appear in the table and figure legends:

CDC	Centers for Disease Control and Prevention
CAN	Certified Nursing Assistant
CI	Confidence Interval
ILI	Influenza Like Illness
LPN	Licensed Practical Nurse
MDS	Minimum Data Set
PPD	Per patient day
RN	Registered Nurse

References

1. Bland JM. Cluster randomised trials in the medical literature: two bibliometric surveys. *BMC Med Res Methodol* 2004;4:21. [PubMed: 15310402]
2. Cornfield J. Randomization by group: a formal analysis. *Am J Epidemiol* 1978; 108(2): 100–102. [PubMed: 707470]
3. Donner A and Donald A. Analysis of data arising from a stratified design with the cluster as unit of randomization. *Stat Med* 1987; 6(1): 43–52. [PubMed: 3576016]
4. Moberg J and Kramer M. A brief history of the cluster randomised trial design. *J R Soc Med* 2015; 108(5): 192–198. [PubMed: 26022551]
5. Dahabreh IJ and Hernán MA. Extending inferences from a randomized trial to a target population. *Eur J Epidemiol* 2019; 34(8): 719–722. [PubMed: 31218483]
6. Susukida R, Crum RM, Ebnesajjad C, et al. Generalizability of findings from randomized controlled trials: application to the National Institute of Drug Abuse Clinical Trials Network. *Addiction* 2017; 112(7): 1210–1219. [PubMed: 28191694]
7. Kennedy-Martin T, Curtis S, Faries D, et al. A literature review on the representativeness of randomized controlled trial samples and implications for the external validity of trial results. *Trials* 2015; 16: 495. [PubMed: 26530985]
8. Rocco MV, Cheung AK, Greene T, et al. The HEMO Study: applicability and generalizability. *Nephrology Dialysis Transplantation* 2004; 20(2): 278–284.
9. Giraudeau B, Caille A, Eldridge SM, et al. Heterogeneity in pragmatic randomised trials: sources and management. *BMC Med* 2022; 20(1): 372. [PubMed: 36303153]
10. Gheorghe A, Roberts TE, Ives JC, et al. Centre selection for clinical trials and the generalisability of results: a mixed methods study. *PLoS One* 2013; 8(2): e56560.

11. Johnson AM, Jones SB, Duncan PW, et al. Hospital recruitment for a pragmatic cluster-randomized clinical trial: Lessons learned from the COMPASS study. *Trials* 2018; 19(1): 74. [PubMed: 29373987]
12. Schwartz D and Lellouch J. Explanatory and Pragmatic Attitudes in Therapeutic Trials. *Journal of Clinical Epidemiology* 2009; 62(5): 499–505. [PubMed: 19348976]
13. Olsen RB, Orr LL, Bell SH, et al. External Validity in Policy Evaluations that Choose Sites Purposively. *J Policy Anal Manage* 2013; 32(1): 107–121. [PubMed: 25152557]
14. Ford I and Norrie J. Pragmatic Trials. *N Engl J Med* 2016; 375(5): 454–463. [PubMed: 27518663]
15. Sox HC and Lewis RJ. Pragmatic Trials: Practical Answers to “Real World” Questions. *JAMA* 2016; 316(11): 1205–1206. [PubMed: 27654606]
16. Gravenstein S, Dahal R, Gozalo PL, et al. A cluster randomized controlled trial comparing relative effectiveness of two licensed influenza vaccines in US nursing homes: Design and rationale. *Clin Trials* 2016; 13(3): 264–274. [PubMed: 26908539]
17. Gravenstein S, Davidson HE, Taljaard M, et al. Comparative effectiveness of high-dose versus standard-dose influenza vaccination on numbers of US nursing home residents admitted to hospital: a cluster-randomised trial. *Lancet Respir Med* 2017; 5(9): 738–746. [PubMed: 28736045]
18. Brown University School of Public Health. Long-Term Care: Facts on Care in the US, <http://ltcfocus.org> (Accessed 20 March 2022).
19. Dahabreh IJ, Haneuse SJ- PA, Robins JM, et al. Study Designs for Extending Causal Inferences From a Randomized Trial to a Target Population. *American Journal of Epidemiology* 2020; 190(8): 1632–1642.
20. Moyo P, Zullo AR, McConeghy KW, et al. Risk factors for pneumonia and influenza hospitalizations in long-term care facility residents: a retrospective cohort study. *BMC Geriatr* 2020; 20(1): 47. [PubMed: 32041538]
21. Childs A, Zullo AR, Joyce NR, et al. The burden of respiratory infections among older adults in long-term care: a systematic review. *BMC Geriatrics* 2019; 19(1): 210. [PubMed: 31382895]
22. Mitchell SL, Teno JM, Kiely DK, et al. The Clinical Course of Advanced Dementia. *N Engl J Med* 2009; 361(16): 1529–1538. [PubMed: 19828530]
23. Mitchell SL, Kiely DK and Hamel MB. Dying with advanced dementia in the nursing home. *Arch Intern Med* 2004; 164(3): 321–326. [PubMed: 14769629]
24. Bosco E, Zullo AR, McConeghy KW, et al. Long-term Care Facility Variation in the Incidence of Pneumonia and Influenza. *Open Forum Infect Dis* 2019; 6(6): ofz230.
25. Austin PC. Using the Standardized Difference to Compare the Prevalence of a Binary Variable Between Two Groups in Observational Research. *Communications in Statistics - Simulation and Computation* 2009; 38(6): 1228–1234.
26. U.S. Food and Drug Administration. Framework for FDA’s Real-World Evidence Program. 2018.
27. Yang R, Carter BL, Gums TH, et al. Selection bias and subject refusal in a cluster-randomized controlled trial. *BMC Medical Research Methodology* 2017; 17(1): 94. [PubMed: 28693427]
28. Eldridge S, Kerry S and Torgerson DJ. Bias in identifying and recruiting participants in cluster randomised trials: what can be done? *BMJ* 2009; 339: b4006. [PubMed: 19819928]
29. Giraudeau B and Ravaud P. Preventing bias in cluster randomised trials. *PLoS Med* 2009; 6(5): e1000065. [PubMed: 19536323]
30. Cole SR and Stuart EA. Generalizing evidence from randomized clinical trials to target populations: The ACTG 320 trial. *Am J Epidemiol* 2010; 172(1): 107–115. [PubMed: 20547574]
31. Dahabreh IJ, Robertson SE, Steingrimsson JA, et al. Extending inferences from a randomized trial to a new target population. *Stat Med* 2020; 39(14): 1999–2014. [PubMed: 32253789]
32. Tipton E and Olsen RB. A Review of Statistical Methods for Generalizing From Evaluations of Educational Interventions. *Educational Researcher* 2018; 47(8): 516–524.
33. Rudolph KE and van der Laan MJ. Robust estimation of encouragement-design intervention effects transported across sites. *J R Stat Soc Series B Stat Methodol* 2017; 79(5): 1509–1525. [PubMed: 29375249]

34. Dahabreh IJ, Robertson SE, Steingrimsson JA, et al. Extending inferences from a cluster randomized trial to a target population. 2022:arXiv:2203.14761. doi:10.48550/arXiv.2203.14761. <https://ui.adsabs.harvard.edu/abs/2022arXiv220314761D> (Accessed 1 March 2022).
35. Robertson SE, Steingrimsson JA and Dahabreh IJ. Analyzing cluster randomized trials designed to support generalizable inferences. 2022:arXiv:2204.02872. doi:10.48550/arXiv.2204.02872. <https://ui.adsabs.harvard.edu/abs/2022arXiv220402872R> (Accessed 1 April 2022).
36. Campbell MK, Piaggio G, Elbourne DR, et al. Consort 2010 statement: extension to cluster randomised trials. *BMJ* 2012; 345: e5661. [PubMed: 22951546]
37. Eldridge S, Ashby D, Bennett C, et al. Internal and external validity of cluster randomised trials: systematic review of recent trials. *BMJ* 2008; 336(7649): 876–880. [PubMed: 18364360]

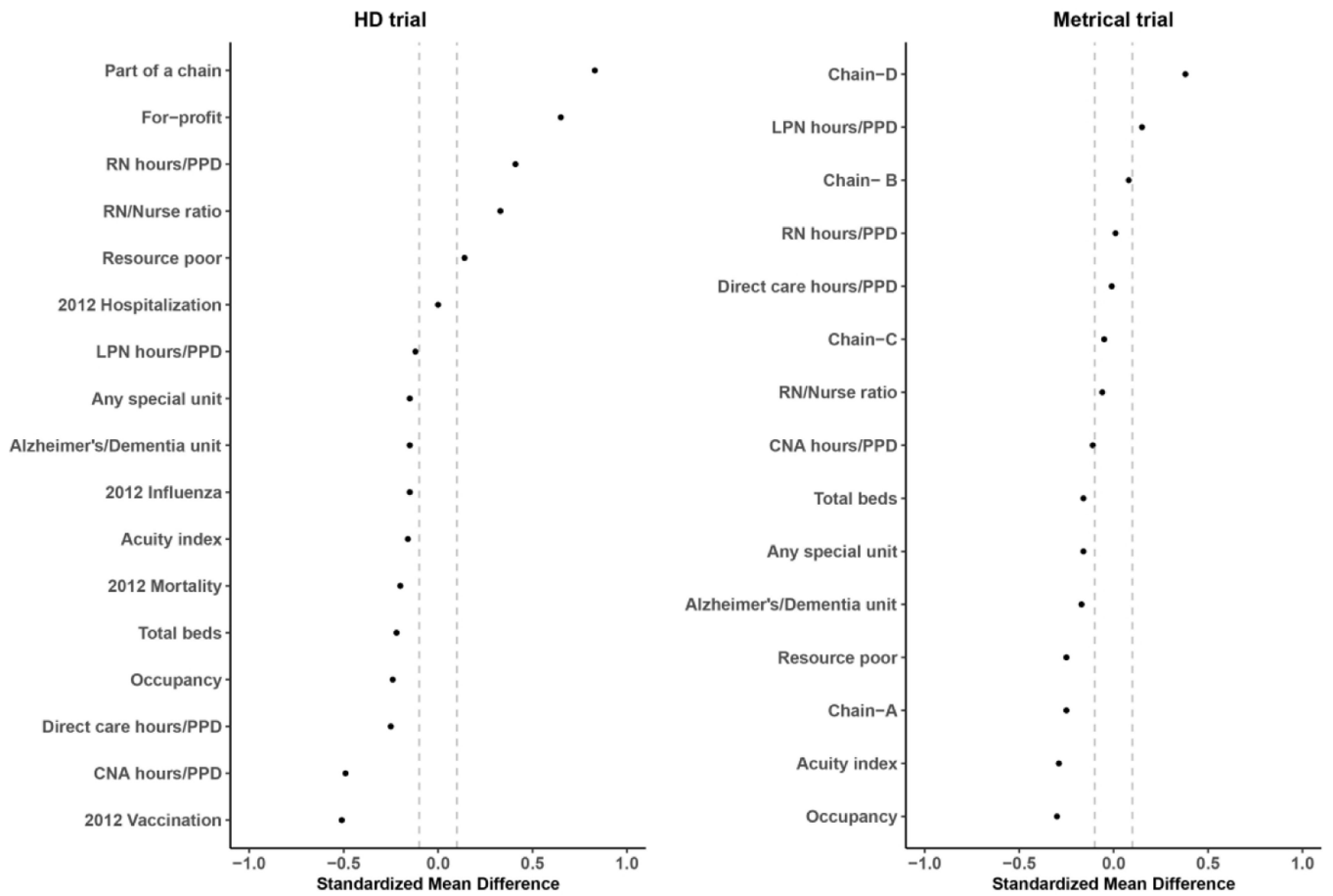


Figure 1. Standardized mean difference of facility level characteristics between randomized and eligible non-randomized nursing homes

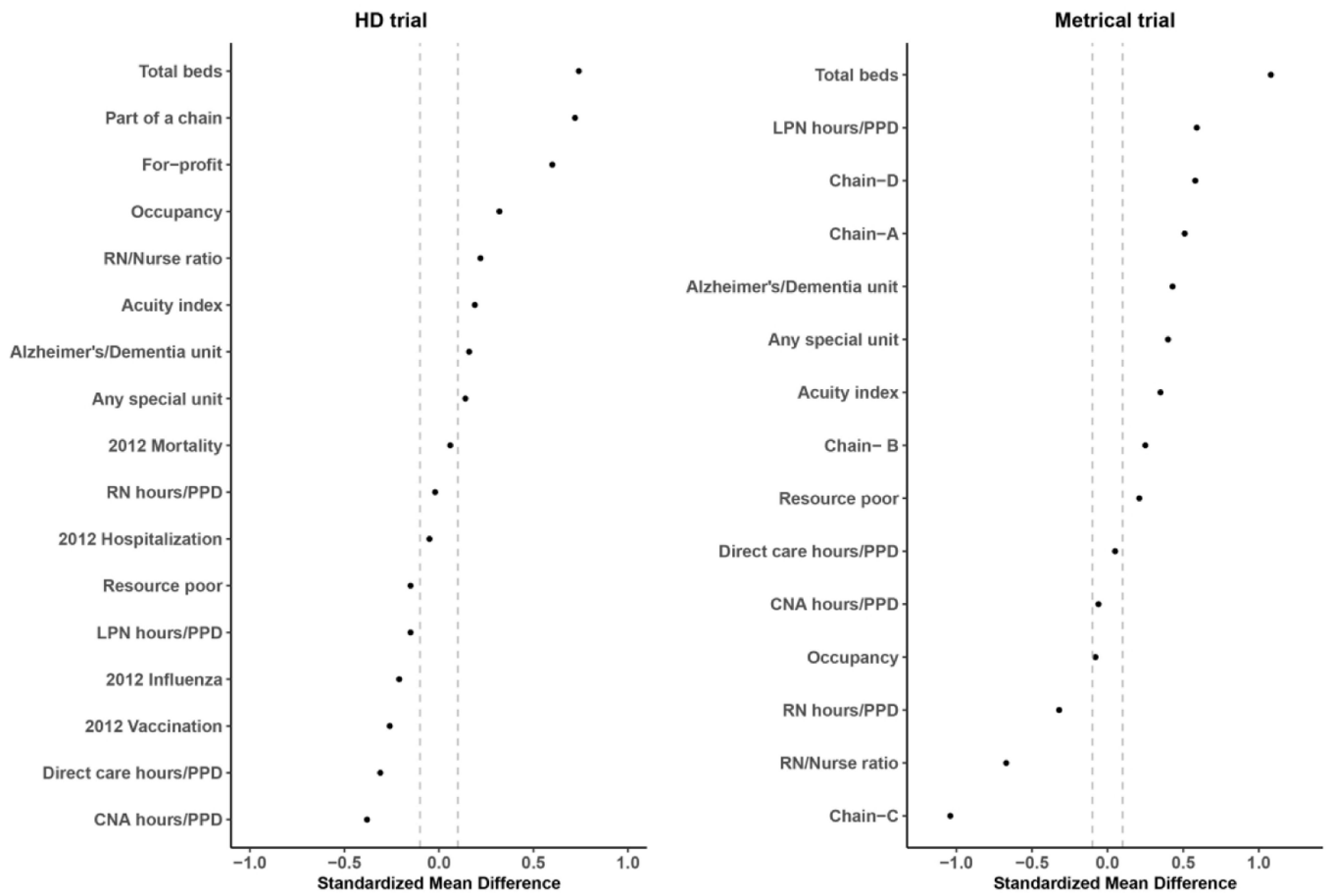


Figure 2. Standardized mean difference of facility level characteristics between randomized and ineligible nursing homes

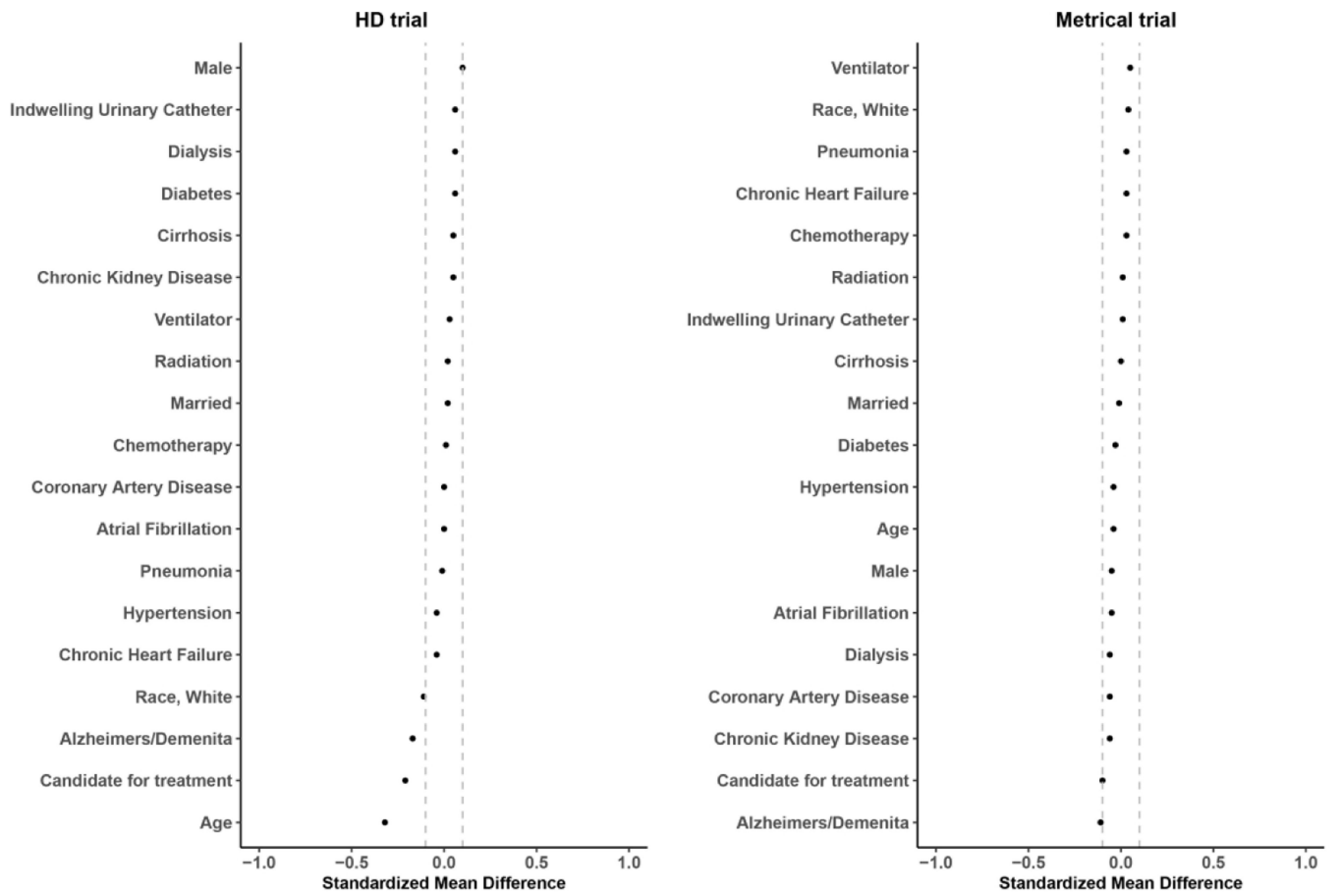


Figure 3. Standardized mean difference of individual level characteristics between individuals residing in randomized and eligible non-randomized nursing homes

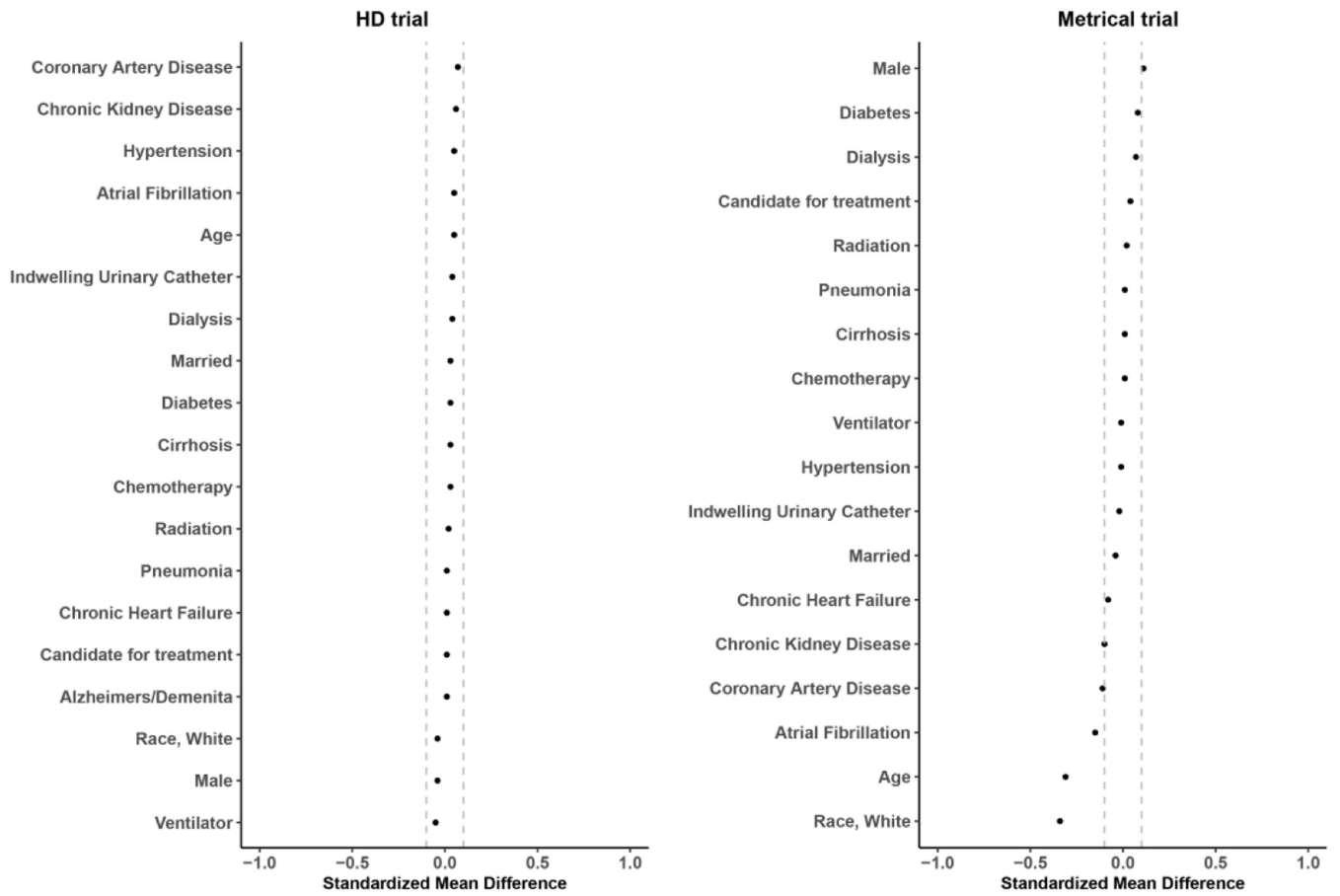


Figure 4. Standardized mean difference of individual level characteristics between individuals residing in randomized and ineligible nursing homes

Table 1.

Comparison of facility and resident characteristics by trial eligibility and participation status in the HD trial

	Randomized facilities	Eligible non-randomized facilities	Ineligible facilities
Facility characteristics (N)	818	3,658	11,026
Total beds (mean/SD)	132.5 (56.0)	145.9 (66.2)	91.9 (53.7)
Resource poor	44 (5.4)	99 (2.7)	1,019 (9.2)
For profit	751 (91.8)	2,443 (66.8)	7,586 (68.8)
Chain facility	702 (85.8)	1,824 (49.9)	6,034 (54.7)
Alzheimer's/Dementia Unit	146 (17.9)	871 (23.8)	1,353 (12.3)
Any special unit	162 (19.8)	947 (25.9)	1,586 (14.4)
RN to nurse ratio (mean/SD)	0.4 (0.2)	0.3 (0.2)	0.4 (0.2)
Direct Care hours/PPD (mean/SD)	3.5 (0.6)	3.6 (0.6)	3.8 (1.4)
RN hours/PPD (mean/SD)	0.5 (0.3)	0.4 (0.3)	0.6 (0.8)
LPN hours/PPD (mean/SD)	0.8 (0.3)	0.8 (0.3)	0.9 (0.7)
CNA hours/PPD (mean/SD)	2.1 (0.4)	2.4 (0.5)	2.4 (1.0)
Acuity index (mean/SD)	12.2 (1.0)	12.4 (1.0)	11.9 (1.9)
Occupancy rate (mean/SD)	84.7 (10.9)	87.3 (10.1)	80.4 (15.6)
2012 mortality rate (mean/SD)	17.2 (5.3)	18.2 (4.7)	16.8 (7.2)
2012 all-cause hospitalization rate (mean/SD)	25.8 (8.2)	25.8 (8.6)	26.3 (11.2)
2012 ILI hospitalization rate (mean/SD)	5.3 (3.0)	5.7 (3.2)	6.1 (4.5)
2012 Influenza vaccination rate (mean/SD)	65.2 (13.9)	72.1 (12.9)	69.6 (19.0)
Resident characteristics (N)	91,179	463,341	806,602
Demographics			
Age (mean/SD)	78.3 (13.7)	82.3 (11.5)	77.6 (14.6)
Male	30977 (34.0)	135,706 (29.3)	287,821 (35.8)
Race, White	68590 (75.2)	369,158 (79.7)	619,102 (77.0)
Married	20,661 (22.7)	101,823 (22.0)	171,438 (21.3)
Candidate for treatment	49,857 (54.7)	300,130 (64.8)	437,424 (54.4)
Chronic conditions			
Alzheimer's/Dementia	45,541 (50.0)	271,384 (58.6)	396,286 (49.3)
Diabetes	32,614 (35.8)	152,571 (32.9)	275,501 (34.2)
Pneumonia	6,455 (7.1)	33,681 (7.3)	55,691 (6.9)
Congestive heart failure	19,445 (21.3)	105,559 (22.8)	169,611 (21.1)
Coronary artery disease	16,297 (17.9)	82,942 (17.9)	122,077 (15.2)
Atrial fibrillation	13,452 (14.8)	68,165 (14.7)	103,879 (12.9)
Hypertension	69,742 (76.5)	361,922 (78.1)	599,241 (74.5)
Cirrhosis	720 (0.8)	1,816 (0.4)	4,223 (0.5)
Chronic kidney disease	9,740 (10.7)	42,580 (9.2)	70,638 (8.8)
Procedures			
Indwelling urinary catheter	8,896 (9.8)	37,426 (8.1)	69,319 (8.6)

	Randomized facilities	Eligible non-randomized facilities	Ineligible facilities
Ventilator	386 (0.4)	1,102 (0.2)	6,647 (0.8)
Chemotherapy	723 (0.8)	3,308 (0.7)	4,516 (0.6)
Radiation therapy	261 (0.3)	793 (0.2)	1,568 (0.2)
Dialysis	2,464 (2.7)	8,197 (1.8)	16,610 (2.1)

CNA = Certified Nursing Assistant; ILI = Influenza Like Illness; LPN = Licensed Practical Nurse; N = Number of Units; PPD = Per Patient Day; RN = Registered Nurse; SD = Standard deviation

^aEntries are frequencies and percent, unless otherwise indicated.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 2.

Comparison of facility and resident characteristics by trial eligibility and participation status in the METRICaL trial

	Randomized facilities	Eligible non-randomized facilities	Ineligible facilities
Facility characteristics (N)^a	80	80	110
Total beds (mean/SD)	103.7 (37.7)	110.5 (45.0)	67.0 (30.0)
Resource poor	8 (10.0)	15 (18.8)	5 (4.6)
Nursing home chain			
A	24 (30.0)	26 (32.5)	84 (76.4)
B	11 (13.8)	9 (11.3)	7 (6.4)
C	15 (18.8)	5 (6.3)	2 (1.8)
D	30 (37.5)	40 (50.0)	17 (15.5)
Alzheimer's/Dementia Unit	19 (23.8)	25 (31.3)	9 (8.2)
Any special unit	21 (26.3)	27 (33.8)	12 (10.9)
RN to nurse ratio (mean/SD)	0.3 (0.2)	0.3 (0.2)	0.5 (0.2)
Direct Care hours/PPD (mean/SD)	3.3 (0.8)	3.3 (0.5)	3.2 (0.8)
RN hours/PPD (mean/SD)	0.4 (0.4)	0.4 (0.3)	0.5 (0.3)
LPN hours/PPD (mean/SD)	0.8 (0.3)	0.8 (0.3)	0.6 (0.3)
CNA hours/PPD (mean/SD)	2.1 (0.5)	2.1 (0.4)	2.1 (0.6)
Acuity index (mean/SD)	12.1(1.2)	12.4 (1.1)	11.7 (1.1)
Occupancy rate (mean/SD)	83.1 (10.6)	86.4 (11.4)	84.0 (12.8)
Resident characteristics (N)^a	6,973	7,431	5,858
Demographics			
Age (mean/SD)	77.6 (13.5)	78.2 (13.1)	81.7 (12.5)
Male	2,605 (37.4)	2,964 (39.9)	1,880 (32.1)
Race, White	4,796 (68.8)	4,978 (67.0)	4,861 (83.0)
Married	1,413 (20.3)	1,550 (20.9)	1,283 (21.9)
Candidate for treatment	2,159 (31.0)	2,642 (35.6)	1,697 (29.0)
Chronic conditions			
Alzheimer's/Dementia	3,002 (43.1)	3,610 (48.6)	2,512 (42.9)
Diabetes	2,310 (33.1)	2,554 (34.4)	1,712 (29.2)
Pneumonia	131 (1.9)	108 (1.5)	103 (1.8)
Congestive heart failure	1,326 (19.0)	1,317 (17.7)	1,312 (22.4)
Coronary artery disease	371 (5.3)	500 (6.7)	470 (8.0)
Atrial fibrillation	359 (5.2)	461 (6.2)	533 (9.1)
Hypertension	4,975 (71.4)	5,447 (73.3)	4,217 (72.0)
Cirrhosis	18 (0.3)	18 (0.2)	13 (0.2)
Chronic kidney disease	310 (4.5)	424 (5.7)	395 (6.7)
Procedures			

	Randomized facilities	Eligible non-randomized facilities	Ineligible facilities
Indwelling urinary catheter	370 (5.3)	367 (4.9)	336 (5.7)
Ventilator ^b	---	---	---
Chemotherapy	42 (0.6)	30 (0.4)	29 (0.5)
Radiation therapy ^b	---	---	---
Dialysis	140 (2.0)	215 (2.9)	67 (1.1)

CNA = Certified Nursing Assistant; ILI = Influenza Like Illness; LPN = Licensed Practical Nurse; N = Number of Units; PPD = Per Patient Day; RN = Registered Nurse; SD = Standard deviation

^aEntries are frequencies and percent, unless otherwise indicated.

^bValues suppressed due to Centers for Medicare and Medicaid Services small cell suppression policy