Published in final edited form as:

Contemp Clin Trials. 2012 November ; 33(6): 1124–1131. doi:10.1016/j.cct.2012.08.004.

Efficacy and effectiveness as aspects of cluster randomized trials with nursing home residents: Methodological insights from a pneumonia prevention trial

Peter H. Van Ness, Ph.D., M.P.H.^a, Peter N. Peduzzi, Ph.D.^b, and Vincent J. Quagliarello, M.D.^c

^aDepartment of Internal Medicine, Geriatrics Section, and the Program on Aging, Yale University School of Medicine

^bDivision of Biostatistics, Yale School of Public Health, and the Yale Center for Analytical Sciences

^cDepartment of Internal Medicine, Infectious Disease Section, Yale University School of Medicine

Abstract

This report discusses how methodological aspects of study efficacy and effectiveness combine in cluster randomized trials in nursing homes. Discussion focuses on the relationships between these study aspects in the Pneumonia Reduction in Institutionalized Disabled Elders (PRIDE) trial, an ongoing cluster randomized clinical trial of pneumonia prevention among nursing home residents launched in October 2009 in Greater New Haven, Connecticut. This clinical trial has enrolled long-term care nursing home residents, over 65 years in age, who have either inadequate oral care or swallowing difficulty, previously identified risk factors for pneumonia. It has used a multicomponent intervention consisting of manual tooth/gum brushing, 0.12% chlorhexidine oral rinse administered twice daily by nurses, and upright feeding positioning at meals to reduce rates of radiographically documented pneumonia. Cluster randomization is attractive for nursing home intervention studies because physical proximity and administrative arrangements make it difficult to deliver different interventions to residents of the same nursing home. Implementing an intervention in an entire home requires integration into the daily life of residents and into the administrative procedures of the nursing home. This characteristic of nursing home cluster randomized trials makes them approximate “real-world” research contexts, but implementation can be challenging. The PRIDE trial of pneumonia prevention utilized specific methodological choices that include both efficacy and effectiveness elements. Cluster randomized trials in nursing homes having elements of both efficacy and effectiveness (i.e., hybrid designs) can address some of the methodological challenges of conducting clinical research in nursing homes; they have distinctive advantages and some limitations.

© 2012 Elsevier Inc. All rights reserved.

Corresponding Author: Peter H. Van Ness, Ph.D., M.P.H., Yale Program on Aging, Yale University School of Medicine, 300 George Street, Suite 775, New Haven, CT 06511-6664 Tel: (203) 737-1958; FAX: (203) 785-4823; peter.vanness@yale.edu Peter N. Peduzzi, Ph.D., Division of Biostatistics, Yale School of Public Health, 60 College Street, LEPH 202, New Haven, CT 06520-8034 Tel: 203-785-5615; FAX: 203-785-6912; peter.peduzzi@yale.edu Vincent J. Quagliarello, M.D., Department of Internal Medicine, Infectious Disease Section, Yale University School of Medicine, 300 Cedar Street, S169, New Haven, CT, 06520-8022 Tel: (203) 785-7571; FAX: (203) 785-3864; vincent.quagliarello@yale.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Drs. Van Ness, Peduzzi, and Quagliarello have no financial or other personal conflicts regarding the composition of this manuscript.

Keywords

Cluster randomized clinical trials; comparative effectiveness research; nursing homes; pneumonia; biostatistics

1. Introduction

The roles of efficacy and effectiveness as criteria for the design and evaluation of clinical biomedical research have been highlighted by the recent appropriation of \$1.1 billion dollars for “Comparative Effectiveness Research” (CER) in the American Recovery and Reinvestment Act of 2009 [1]. This renewed commitment to effectiveness research has intensified discussion of this topic in policy reports [2,3], journal articles [4-7], and academic conferences [8,9]. Considerations of efficacy and effectiveness have figured prominently in the thinking of designers and critics of clinical trials ever since Schwartz and Lellouch [10] distinguished between pragmatic and explanatory attitudes in trial design. The pragmatic attitude is interested in discriminating between options for care and is associated with trials in “real-world” situations that compare the *effectiveness* of several active treatments, including usual care but not placebo comparison groups. The explanatory attitude is most interested in testing causal hypotheses and is associated with trials in idealized conditions that compare the *efficacy* of several treatments (e.g., placebo-controlled drug trials). CER focuses on effectiveness studies that often have large sample sizes, heterogeneous patient samples, long-term follow-ups, and patient-centered outcomes (e.g., morbidity, mortality, and quality of life). A major challenge for CER investigation is to enhance effectiveness, while still maintaining the internal validity of the trial, and without incurring inordinate costs.

The Federal Coordinating Council for Comparative Effectiveness Research reports that “Comparative effectiveness differs from efficacy research because it is ultimately applicable to ‘real-world’ needs and decisions faced by patients, clinicians, and other decision makers” [3]. Proponents of CER have advocated the use of clustered randomized designs [11,12] because they often approximate “real-world” research contexts. In the nursing home setting, physical proximity and administrative arrangements make it difficult to deliver different interventions to persons residing in the same nursing home; the use of a clustered design is therefore advantageous. Although many types of clinical biomedical research are hybrid combinations of both efficacy and effectiveness elements, [13-15] clustered designs combine these elements in distinctive ways. In this report, we discuss how efficacy and effectiveness considerations have played a role in the design and conduct of the PRIDE trial, an ongoing cluster randomized trial of an enhanced oral hygiene intervention to prevent pneumonia in nursing home residents. We also discuss methodological responses to the challenges of conducting clinical research in nursing home settings.

2. Methods

2.1. Overview of the PRIDE Trial

PRIDE was based on the premise that a multicomponent intervention, which targets known modifiable risk factors, can reduce nursing home acquired pneumonia. As an initial step, an observational study was conducted to identify modifiable risk factors for nursing home-acquired pneumonia [16]. It identified and validated inadequate oral care and swallowing difficulty as modifiable risk factors independently associated with pneumonia among nursing home residents. In a subsequent pilot study, a multicomponent intervention (i.e., manual tooth/gum brushing + 0.12% chlorhexidine oral rinse twice daily, plus upright positioning during meals) was tested and yielded evidence for high feasibility within the

nursing home setting, high adherence by nursing home staff, and effectiveness in risk factor reduction (i.e., improving oral hygiene and swallowing) [17]. These studies provided empirical evidence justifying the conduct of PRIDE to evaluate if the multicomponent intervention could reduce pneumonia in nursing home residents.

2.2 Design of the PRIDE Trial

PRIDE was designed as a cluster randomized trial to test the superiority of the multicomponent intervention in reducing the rate of pneumonia in older nursing home residents, compared to usual care (i.e., oral hygiene and feeding positioning) in control nursing homes. A schematic representation of the PRIDE trial design, in which nursing homes (i.e., clusters) are randomized, is provided in Figure 1a; a contrasting schematic representation for a conventional parallel group design in which individual participants are randomized is provided in Figure 1b. The primary outcome of the PRIDE trial is time to first pneumonia. Although randomization is being carried out at the level of nursing homes, the analysis is being conducted at the level of individual residents controlling for the clustering effect. A clustered design, in which nursing homes instead of individuals are the unit of randomization, was used to avoid potential for contamination of controls when randomizing at the level of the participant or nursing unit. Contamination in the control arm would compromise the ability of the study to detect effects because treatments would become more similar among intervention and control groups. Other reasons for using a clustered design included greater efficiency in staff training, improved feasibility of applying the intervention facility-wide, and ready accommodation of usual care in the control group [18]. The main disadvantage with this design is the larger sample size required to account for the correlation among residents within a home (i.e., intracluster correlation). However, efficient procedures made possible by providing all residents of a nursing home with the same intervention might mitigate the added costs of a larger sample size. Another potential disadvantage, when there are few homes, is the greater likelihood for imbalance in baseline characteristics between residents in intervention and control homes owing to differences in patient mix among the homes. To improve the likelihood that the intervention and control arms will have similar characteristics, participating homes in the PRIDE trial were stratified based on a variable that is related to the delivery of the intervention. Thirty-six homes were stratified into two strata based on the average number of minutes per day that nurses' aides spend with residents as reported in the Center for Medicare and Medicaid Services Quality Measure database (i.e., < 140 vs. 140 minutes).

The setting for the trial is Connecticut-area nursing homes providing skilled nursing care and certified for Medicare or Medicaid. The participants in the trial are long-term care nursing home residents over 65 years in age who provided informed consent either individually or by personally identified proxy. To be eligible for enrollment, residents must have resided in the home for at least four weeks and must have at least one of the two previously identified risk factors for pneumonia: inadequate oral care indicated by a plaque score ≥ 1 and swallowing difficulty as identified by nursing home staff. The study's target sample size of 828 study participants accounts for the clustering design effect. The design effect is an inflation factor to sample size because the clustering of residents within a nursing home introduces a correlation in which residents within a home have a greater likelihood of being similar to one another than to residents from another home. The design effect is a function of the intracluster correlation and the number of residents to be enrolled per nursing home. Based on data available at the time of study planning, the estimated design effect required inflating the sample size in PRIDE by 21.5%. Using an intracluster correlation coefficient (ICC) of 0.005 from a previous observational study conducted by us [16], a sample of 828 study participants provides 80% power to detect, with type I error of

0.05 (2-sided), a hypothesized 25% reduction in the 2.5 year cumulative pneumonia rate with intervention relative to control.

In the homes randomized to the intervention group, the nursing home staff administers the intervention for the duration of the trial, after training and with continued surveillance by study investigators. In homes randomized to the control arm, usual care is provided to study participants. The planned study duration is 4 years with 2.5 years for enrollment overlapping with 2.5 years of surveillance for the development of radiographically documented pneumonia. Requests for chest x-rays are provided by nursing home physicians; final adjudication of pneumonia is made by Infectious Disease physician investigators, blinded to intervention or control status of participants. The study has been monitored by an independent Data and Safety Monitoring Board. PRIDE was launched in October 2009.

To ensure an adequate sample size for PRIDE, screening for enrollment took place in two phases. Initially, all residents in a home were screened over a two-month period. Once the initial screening and consenting had been completed, nursing homes were randomized to intervention or control arms (i.e., prevalent enrollment). All eligible and consented residents in a home then received the treatment assigned to the home. Thereafter, at subsequent three-month intervals, new and initially ineligible residents were screened for eligibility and enrolled for up to 2.5 years (i.e., incident enrollment).

This strategy made it possible to recruit a larger number of residents from each nursing home and reduce the number of nursing homes required to achieve the target sample size. Enrolling residents on an incident basis, however, increased the risk of selection bias because homes had already been randomized when such enrollments occurred and there was a risk that incident enrollees (or their proxies) would know whether their nursing home was an intervention or control home. In an attempt to prevent unmasking of potential incident enrollees (and their proxies), all public information available about PRIDE in each home (e.g., Study Brochures, Study Posters) described the study without revealing the randomization assignment.

3. Results

3.1 Efficacy and Effectiveness as Aspects of PRIDE

Because of practical and feasibility considerations, most clinical trials include aspects of both efficacy and effectiveness (i.e., they are hybrid designs). In general, efficacy relates to internal validity and effectiveness to external validity. Comprehensive reviews of the distinctions between efficacy and effectiveness clinical trials have been reported [19,20]. Table 1 provides a summary of some key attributes characteristic of randomized clinical effectiveness trials along with their advantages and disadvantages. The attributes include study participants, setting, study duration, sample size, comparison groups, treatment modality, health outcome type, adherence monitoring, and intent-to-treat analyses.

3.2.1. Participants—PRIDE study participants were restricted to nursing home residents with at least one of the two modifiable risk factors because the multicomponent intervention was designed to target these risk factors. While this strategy preserves the study's internal validity to detect treatment effects, it could potentially limit its external validity since results will be applicable only to a population of nursing home residents with impaired oral hygiene and/ or swallowing difficulty. However, PRIDE recruitment data have demonstrated that most long-term older residents (i.e., age >65 years) at each nursing home have at least one of the two risk factors, making external validity more achievable.

3.2.2. Setting—PRIDE has been conducted in community-based nursing homes rather than within specialized research institutions or academic medical centers. The nursing home environment is very dynamic with residents, staff, administrative leadership, and treatment regimens frequently changing [21]. Residents drop out of studies due to frailty and mortality from underlying comorbid conditions. Administrators, medical directors, nursing directors and staff have high rates of turnover due to fiscal instability, inadequate staffing, and regulatory burdens. Administrative procedures, and, consequently, study protocols, can be interrupted by state audits, receivership of homes for financial issues, and infectious disease outbreaks unrelated to study outcomes. Nursing homes that are members of corporate chains can have less flexible corporate-wide policies that become foci of labor conflict; independent nonprofit homes can be subject to resource limitations that inhibit consistent implementation of study protocols. Conducting a cluster randomized clinical trial in a nursing home with these dynamic features promotes study effectiveness in such a real-world setting, but it can constrain study efficacy by reducing the feasibility of completing the trial.

3.2.3 Study Duration—Participants in PRIDE are being followed for the development of pneumonia for 2.5 years (i.e., an average duration of time that a long-term care resident is housed within a nursing home). This targeted duration has been limited because of the expected high annual censoring rate due to mortality (17%) and transfers out of the nursing home (3%) [16]. However, because of the high censoring and pneumonia rates, longer follow-up would yield few additional cases of pneumonia and not be cost-effective. Thus, the follow-up duration is relatively long-term for this population, which is more consistent with an effectiveness design.

3.2.4 Sample Size—Considering the difficulty of recruiting and retaining study participants in cluster trials, in nursing homes [21] and other group contexts [22], the target sample size of 828 participants can be considered large for a nursing home study. Ideally, in a CER study, sample size would be calculated to accommodate not only an effect in the entire study sample but also in subgroups of special interest (e.g., sociodemographic subgroups). Because PRIDE was restricted to the Connecticut geographic area due to fiscal and feasibility considerations, the size of the study was limited by the number of available nursing homes. Thus, there may not be adequate power to detect treatment effects in key subgroups of participants (e.g., race and gender). This geographic restriction also prevents the application of study findings to a wider population of nursing homes. Therefore, the sample size could be considered intermediate between an efficacy and effectiveness trial.

3.2.5. Comparison Group—The comparison group for PRIDE is usual care (i.e., the usual standard of oral care and feeding positioning that each nursing home utilized). This type of comparison group is typically used in effectiveness research. An advantage of the clustered design is that entire nursing homes are assigned to intervention or control groups, thereby allowing them to be administered as consistently and efficiently as they might be in a realistic nursing home setting.

3.2.6. Treatment Modality—The PRIDE intervention was designed to target risk factors associated with nursing home acquired pneumonia. Nursing home staff members, upon receiving comprehensive training by research personnel, adopt the multicomponent intervention (i.e., oral brushing with toothpaste [including teeth, gums and/or dentures], topical oral chlorhexidine administered by nurses, and upright positioning during meals) during their daily care plan of the participants. Adopting this protocol adds approximately 10-15 minutes of resident care per day. Thus, the treatment modality resembles a real-world circumstance to a greater degree than it would if the intervention were administered by trained research personnel. The clustered design of the PRIDE trial facilitated this treatment

modality by allowing staff training and intervention implementation to be performed efficiently throughout an entire intervention nursing home. However, adherence to the intervention is being monitored in the intervention homes, as it would be in an efficacy study. Thus, the treatment modality has elements of both effectiveness and efficacy.

3.2.7. Health Outcome Type—PRIDE is using a patient-centered outcome of a radiographically documented pneumonia, which is based on both clinical features and chest x-ray findings [16]. Using an outcome that is important to both patients and providers is characteristic of effectiveness trials. However, the primary outcome is being adjudicated by a committee of investigators blinded to participant treatment assignment instead of by individual physicians treating residents in the enrolled study nursing homes. PRIDE adjudications were based on both radiographic evidence of pneumonia and at least two compatible clinical features (e.g., cough, fever, dyspnea). Such central adjudication is common in efficacy but not effectiveness trials.

3.2.8. Adherence Monitoring—In PRIDE, adherence to the delivery of the intervention by the nursing home staff is carefully monitored, and these data are used to identify areas for adherence improvement. Surveillance for adherence to the intervention is conducted by study personnel during unannounced observations twice per month over the course of follow-up. Adherence is determined by means of direct observation of staff administering the intervention, and by documentation of expected vs. observed toothpaste tube weights and chlorhexidine residual volumes over a previous two-week period. Toothpaste and chlorhexidine use is not monitored in the usual care control group. Although adherence monitoring is not typically done in an effectiveness study, it was important in PRIDE to maintain the integrity of the intervention arm to preserve the power of the study.

3.2.9. Intent-To-Treat Analysis—A distinctive attribute of effectiveness studies is analysis by intent-to-treat (ITT) in which all participants are analyzed according to the study arm to which they were randomly assigned, regardless of the extent to which they adhered to the intervention. ITT mimics what happens in the real world and usually gives a conservative estimate of the treatment effect in a superiority trial. This is the approach being used for the analysis of PRIDE. In contrast, many efficacy studies, such as those typically conducted by the pharmaceutical industry to register new drugs, include a per-protocol analysis which includes only those participants who adhered to regimens described in the study protocol in attempt to estimate the “true effect” of treatment. Analyzing data only from participants who fully adhere to the intervention, or according to their degree of adherence, can lead to bias in an outcome analysis if reasons for nonadherence differ between intervention and control groups, which often happens in practice [23]. Recent statistical advances permit the analysis of treatment effects in the presence of nonadherence [24] [25]. These methods, however, cannot be applied in PRIDE because adherence data are only being collected in the intervention nursing homes. These data will be used to examine the impact of nonadherence in the intervention arm, (e.g., dose-response effect).

4. Conclusions

An objective of CER is to conduct effectiveness studies that are generalizable to the real world. Conducting such studies in the nursing home setting is challenging because the nursing home environment is very dynamic. Although nursing homes provide a real-world setting, there are practical challenges to conducting effectiveness studies in this environment. Because of these challenges, PRIDE incorporated elements of both efficacy and effectiveness in a hybrid design. Table 2 summarizes several methodological challenges that arose in the design of PRIDE and how these challenges were addressed.

The choice of a clustered design for the PRIDE trial was motivated primarily by a concern to avoid contamination of intervention and control groups—an efficacy concern. This choice also provided the opportunity to integrate the administration of the intervention into the regular care of nursing home resident by having healthcare providers, instead of research staff, administer the brushing of teeth and dispensing of topical oral chlorhexidine. This strategy enhanced generalizability and real-world aspects of the study (i.e., its effectiveness aspect).

Clustered designs present opportunities for increased effectiveness but they also introduce characteristic constraints. A statistical constraint of clustered designs is the need to account for the correlation among subjects within a cluster, which requires inflating the sample size. Often an estimate of the intraclass correlation is not available at the time of study planning, making sample size determination uncertain and potentially leading to an underpowered study. The PRIDE investigators collected pilot data to obtain an estimate of the intraclass correlation. Interim monitoring of the ICC is being done by the PRIDE DSMB to determine whether sample size readjustment is needed to maintain the power of the study. The cluster design effect also placed additional constraints on the study because the number of available nursing homes was limited by financial resources.

In summary, there is a need to develop and test therapies for older persons, particularly in real-world settings like nursing homes. Clustered designs offer one approach, involving characteristic opportunities and constraints. Geriatric and gerontologic researchers have considerable experience and expertise working with older study participants and grappling with the distinctive clinical [21] and statistical [26] challenges they pose. The PRIDE trial is contributing to current efforts to advance comparative effectiveness research among nursing home residents.

Acknowledgments

All co-authors contributed text and made editorial contributions to the manuscript. Dr. Van Ness wrote the initial draft of the manuscript, integrated co-author contributions into the evolving manuscript, and prepared the manuscript for submission. The first author thanks Heather Allore for her helpful comments.

Dr. Van Ness was supported in part by funding from the Claude D. Pepper Older Americans Independence Center at the Yale University School of Medicine (P30AG021342). Dr. Peduzzi was supported in part by CTSA Grant Number UL1 RR024139 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH roadmap for Medical Research. Dr. Quagliarello was supported in part by grants (R01 AG030575 and K07 AG030093) from the National Institute on Aging (NIA). The contents of this article are solely the responsibility of the authors and do not necessarily represent the official view of the NIH or NCRR. The organizations funding this essay had no role in the preparation, review, or approval of this manuscript.

References

1. Steinbrook R. Health care and the American Recovery and Reinvestment Act. *N Engl J Med*. 2009; 360:1057–1060. [PubMed: 19224738]
2. IOM (Institute of Medicine). Initial National Priorities for Comparative Effectiveness Research. National Academic Press; Washington, DC: 2009.
3. Federal Coordinating Council for Comparative Effectiveness Research. [August 3, 2011] Report to the President and the Congress. 2009. <http://www.hhs.gov/recovery/programs/cer/cerannualrpt.pdf>.
4. Conway PH, Clancy C. Comparative-Effectiveness Research: Implications of the Federal Coordinating Council's Report. *N Engl J Med*. 2009; 361:328–330. [PubMed: 19567829]
5. Iglehart JK. Prioritizing comparative-effectiveness research: IOM recommendations. *N Engl J Med*. 2009; 361:325–328. [PubMed: 19567828]
6. Lauer MS, Collins FS. Using science to improve the nation's health system: NIH's commitment to comparative effectiveness research. *JAMA*. 2010; 303:2182–2183. [PubMed: 20516419]

7. Sox HC, Greenfield S. Comparative effectiveness research: A report from the Institute of Medicine. *Ann Intern Med.* 2009; 151:203–205. [PubMed: 19567618]
8. Statistical Issues in Comparative Effectiveness Research (CER). Paper presented at: University of Pennsylvania Annual Conference on Statistical Issues in Clinical Trials: From Bench to Bedside to Community; University of Pennsylvania School of Medicine, Philadelphia, PA. 2010;
9. Clinical and Comparative Effectiveness Research Methods II. Paper presented at: Agency for Healthcare Research and Quality Symposium; Agency for Healthcare Research and Quality Conference Center, Rockville, MD. 2009;
10. Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in therapeutical trials. *J Chronic Dis.* 1967; 20:637–648. [PubMed: 4860352]
11. Hayes, RJ.; Moulton, LH. Cluster randomised trials. Chapman & Hall/ CRC; Boca Raton, FL: 2009.
12. Donner, A.; Klar, N. Design and analysis of cluster randomization trials in health research. Arnold Publishers; London: 2000.
13. Atkins MS, Frazier SL, Cappella E. Hybrid research models: Natural opportunities for examining mental health in context. *Clin Psychol Sci Pract.* 2006; 13:105–108.
14. Carroll KM, Rounsaville BJ. Bridging the gap: A hybrid model to link efficacy and effectiveness research in substance abuse treatment. *Psychiatr Serv.* 2003; 54:333–339. [PubMed: 12610240]
15. Armitage P. Attitudes in clinical trials. *Stat Med.* 1998; 17:2675–2683. [PubMed: 9881414]
16. Quagliarello V, Ginter S, Han L, Van Ness P, Allore H, Tinetti M. Modifiable risk factors for nursing home-acquired pneumonia. *Clin Infect Dis.* 2005; 40:1–6. [PubMed: 15614684]
17. Quagliarello V, Juthani-Mehta M, Ginter S, Towle V, Allore H, Tinetti M. Pilot testing of intervention protocols to prevent pneumonia in nursing home residents. *J Am Geriatr Soc.* 2009; 57:1226–1231. [PubMed: 19558483]
18. Margolis KL, Kerby TJ, Asche SE, Bergdall AR, Maciosek MV, O'Connor PJ, Sperl-Hillen JM. Design and rationale for Home Blood Pressure Telemonitoring and Case Management to Control Hypertension (HyperLink): a cluster randomized trial. *Contemp Clin Trials.* 2012
19. Gartlehner G, Hansen RA, Nissman D, Lohr KN, Carey TS. A simple and valid tool distinguished efficacy from effectiveness studies. *J Clin Epidemiol.* 2006; 59:1040–1048. [PubMed: 16980143]
20. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, Tunis S, Bergel E, Harvey I, Magid DJ, Chalkidou K. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol.* 2009; 62:464–475. [PubMed: 19348971]
21. Juthani-Mehta M, Quagliarello VJ. Infectious diseases in the nursing home setting: challenges and opportunities for clinical investigation. *Clin Infect Dis.* 2010; 51:931–936. [PubMed: 20822459]
22. Andrae SJ, Halanych JH, Cherrington A, Safford MM. Recruitment of a rural, southern, predominantly African-American population into a diabetes self-management trial. *Contemp Clin Trials.* 2012; 33:499–506. [PubMed: 22349456]
23. Friedman, LM.; Furberg, CD.; DeMets, DL. Fundamentals of clinical trials. Springer; New York: 1998.
24. Jin H, Rubin DB. Principal stratification for causal inference with extended partial compliance. *J Am Stat Assoc.* 2008; 103:101–111.
25. Robins JM, Finklestein DM. Correcting for noncompliance and dependent censoring in an AIDS clinical trial with inverse probability of censoring weighted (IPCW) log-rank tests. *Biometrics.* 2000; 56:779–788. [PubMed: 10985216]
26. Van Ness PH, Charpentier PA, Ip EH, Leng I, Murphy TE, Tooze JA, Allore HG. Gerontologic biostatistics: The statistical challenges of clinical research with older study participants. *J Am Geriatr Soc.* 2010; 58:1386–1392. PMID 20533963. [PubMed: 20533963]

Figure 1a.
Schematic Representation of the PRIDE Trial Design

Figure 1b.
Schematic Representation of a Conventional Parallel Group Randomized Clinical Trial Design

Table 1**Attributes of Clinical Effectiveness Trials: Advantages and Disadvantages**

	Trial Attribute	Characteristic of Effectiveness Trial	Advantages (A) and Disadvantages (D)
1	Participants	The eligibility criteria are not restrictive.	(A)The study sample is more heterogeneous, reflecting real-world populations; (D) heterogeneity increases variability thereby requiring a larger study sample size.
2	Setting	The setting is realistic, being a hospital, nursing home or other context in which participants reside.	(A) The results yielded by the study are more generalizable to physicians, patients, and others who make treatment decisions in real world settings; (D) it is more difficult to implement study protocols in real-world settings because of the dynamic nature of the environments.
3	Study Duration	The duration is longer.	(A) Longer follow-up generates more outcomes to detect important but modest treatment effects in heterogeneous populations and to examine the durability of effects; (D) costs to conduct the trial are increased and losses to follow-up are more likely.
4	Sample Size	The sample size is large.	(A) Larger sample size makes it possible to examine effects in important subgroups; (D) costs are increased.
5	Comparison Group	The control group receives an active treatment or usual care, instead of a placebo or another type of inactive treatment.	(A) Treatments received by control group participants are more likely to be beneficial and are more ethically justifiable; (D) use of active controls makes it harder to, requiring a larger sample size.
6	Treatment Modality	The treatment is administered by healthcare providers and not by research specialists.	(A) It is more probable that a successful intervention will be employed in real-world treatment practice; (D) it is more likely that intervention administration will deviate from protocol.
7	Health Outcome Type	The outcome is patient-centered, such as mortality or morbidity rather than a biomarker or other surrogate outcome.	(A) Study results are more meaningful to patients, providers and decision-makers; (D) larger sample sizes are required to study patient-centered outcomes than intermediate outcomes.
8	Adherence Monitoring	The monitoring of adherence is unobtrusive or nonexistent, and little or no efforts are made to ensure adherence.	(A) Less time and money need to be devoted to monitoring adherence; (D) nonadherence can dilute the treatment benefit in the intervention arm and requires larger sample size.
9	Intent-to-Treat Analysis	Intent-to-treat analysis is employed.	(A) The analysis reflects real-world circumstances in which patients do not always adhere to their prescribed treatment assignment; (D) in the presence of nonadherence the efficacy of an intervention cannot be accurately estimated.

Table 2**Methodological Challenges and Responses for Clinical Trials of Nursing Home Residents: Implications for the PRIDE Study**

Older-Person Life Factor	Methodological Challenge	Methodological Responses	Implication
Residential clustering (e.g., in nursing homes) as an intervention implementation issue	Potential contamination of the control group (e.g., with nursing home residents or staff adopting elements of the intervention they observe in nearby residents)	Randomization of nursing homes, instead of individual residents, to intervention or control groups	Randomization of homes instead of individuals promotes effectiveness by allowing the intervention to be administered by trained healthcare providers (e.g., nursing home staff).
Residential clustering as a sample size and estimation issue	The clustering of study participants introduces a correlation among individuals in a cluster (i.e., intracluster correlation).	The study sample size needs to be inflated to account for the intracluster correlation and the analysis requires a robust variance estimator to ensure that standard errors are not artificially reduced by clustering.	These responses ensure adequate study power and proper test statistics.
Residential clustering as a randomization issue	The randomization of nursing homes rather than individuals is more likely to result in imbalances on important baseline characteristics between intervention and control groups because the unit of randomization is the home instead of the individual (i.e., there may be differences in case mix between homes).	The analytical plan includes a pre-specified set of variables for covariate adjustment.	Covariate adjustment for baseline imbalances controls for systematic bias, providing a fair comparison of treatments.
Nursing home usual care varies in practice, with oral hygiene care provided at different levels in different nursing homes	Heterogeneity in the delivery of usual care among control nursing homes	Stratified randomization (i.e., randomization occurring within strata defined by the daily time period that nursing home staff members devote to residents)	This approach may not completely account for all heterogeneity in the control group but it avoids an unrealistic standardization of usual care practice that would diminish the effectiveness aspect of the trial.
Disinclination or inability to participate in research, reflecting the seriously ill status of some older persons in nursing homes	Lower than expected rates of recruitment, with proxy consent for individual study participants required in the majority of eligible residents	Recruitment of incident residents in which research staff return to nursing homes after prevalent enrollment ends to enroll new residents, or enroll residents that have recently satisfied study eligibility requirements	This response promotes effectiveness by having a higher proportion of nursing home residents enrolled in the study; it risks some selection bias because incident enrollees may know the intervention status of their home.
Multiple morbidities reflecting the seriously ill status of older persons lead to difficulties in administering the intervention	Maintaining separation of intervention and control groups	Monitoring adherence to the study protocol in the intervention homes	This response promotes efficacy by assuring adherence to the study protocol by persons in the intervention arm.
Older persons in nursing homes receive medical care from nursing home physicians who are responsible for ordering tests such as chest x-rays	The primary outcome of radiographically documented pneumonia might be underreported if chest x-rays are not available	Identification of a secondary outcome (i.e., clinical lower respiratory tract infection) that requires compatible clinical features, but does not require a chest x-ray	Requirement of radiographically documented pneumonia promotes efficacy by avoiding false positive outcomes; the secondary outcome reflects real-world circumstances in which the best diagnostic information is not always available.

Older-Person Life Factor	Methodological Challenge	Methodological Responses	Implication
Nursing home patients reside in institutions with different and changing administrative and economic characteristics (e.g., corporate chain versus independent nonprofit ownership)	Financial constraints can force nursing homes to close during the course of the study; labor strife can prevent intervention protocols from being implemented	The primary analyses is a time-to-event analysis in which study participants in nursing homes that close are censored at the time of closing, making use of all available surveillance data; every effort is made to continue outcome surveillance even if the intervention can no longer be administered.	Administrative or economic decisions that result in study participant withdrawal inhibit the ability to determine the true effect (efficacy) of an intervention.